

Not the last word on the BSE crisis

The publication of apparently conclusive evidence linking BSE in cattle to CJD in humans strengthens the case for a full review of how the BSE issue was handled. Britain would not be the only beneficiary.

If politics was an exact science, then the two papers published in this week's issue on the relationship between bovine spongiform encephalopathy (BSE) in cattle and Creutzfeldt-Jakob disease (CJD) in humans might be read as the final chapter — if not the post-script — in an unhappy saga which opened almost exactly ten years ago. Both papers, one based on work carried out by a team headed by Moira Bruce at the Institute of Animal Health in Edinburgh, the other by researchers led by John Collinge at St Mary's Hospital in London, point in the same direction, namely that the agent responsible for the two diseases is identical, and that the human disease was contracted from eating contaminated beef (see pages 498–501, 448–450 and 437–438). And both will therefore help to dispel whatever few doubts may continue to linger that alternative explanations remain possible.

But if the scientific answer to the question “can BSE cause CJD?” now appears to have been answered, many political questions are still unresolved. How did the government and its advisers react to the initial news of a possible new type of disease spreading through British herds? Was the appropriate balance struck between the need to warn consumers and a concern not to undermine a multibillion pound cattle industry? And were official actions in any way to blame for the 20 new-variant-CJD deaths that have occurred so far — an issue of critical importance to the relatives of those who have died, in their fight to recover the costs of providing medical treatment?

There are deeper issues that reach to the heart of the way in which government-funded science is organized, both in Britain and elsewhere. Is there a genuine risk that government research institutes and advisory committees may be influenced by the mission statements of the departments to which they are responsible in a way that distorts their assessment of sensitive issues? Does the BSE crisis strengthen the

argument for more investment in basic long-term research, not linked to a specific strategic objective, at such institutes? Or does it give weight to the alternative argument that such work should be done by disinterested outside groups, for example in universities?

Of course, many important lessons have already been learnt in Whitehall. There is, for example, reported to be a far greater awareness of the need to listen to external scientific experts and to make research data available to their scrutiny — even if this could have embarrassing implications. The government's long-awaited reaction to the proposals for a Food Standards Agency is likely to reflect an awareness of the need to make consumer interests more broadly represented in food policy and food research. Closer attention is being paid to public statements about what ministers consider to be ‘safe’ or ‘dangerous’. And even the new research provides no reason for believing that health measures taken in recent years to curb the spread of BSE have been inadequate.

But it is important that the issue is not left there. There is also a question of how to build public trust in the way in which the government handles scientific issues in a modern, media-based society — an issue of broad interest to any industrial democracy. There is no turning back the clock, and it is important that such a review does not turn into a witch-hunt; many of those who have found themselves in the dock in recent years undoubtedly acted in what they considered to be the best public interest while struggling with incomplete scientific information. But there are still lessons to be learnt, and further changes to be made, to reduce the chances of such a major health hazard recurring. A review that keeps this goal clearly in mind, gathers the appropriate evidence and delivers a clear and impartial judgement on why mistakes were made and what can be done to avoid them, would be a welcome and refreshing step. □

Time for imagination in Prague

The Czech government should not miss an opportunity to give its scientists a needed shot in the arm.

The Czech Republic inevitably has to pay for the acts of God that befall it, the most recent being this summer's devastating floods. But it should try to leave science untouched. As in all former communist countries facing economic difficulties, science is in danger of growing old and dying out. Young people contemplating a career in research are discouraged both by low wages and by the lack of any indication that Czech science has a robust future.

In seeking ways to pay for the flood damage, the Czech government should ensure that it does not worsen the prospects for its scientific future (see page 430). Scientists should be exempt from any wage cuts, and freed from the wage freeze recently imposed on all civil servants. The government should also do what it can to ensure that bold plans by the Academy of Sciences to develop a high-power laser donated by the Max Planck Institute for Quantum Optics in Garching, near Munich, into an inter-

national research facility are not put at risk.

The laser project offers Czech scientists a unique opportunity to demonstrate to Western colleagues their ability to operate the type of large-scale facility that is common in the European Union. At the same time, it will enable young researchers to play host to European guest scientists, and thus become fully integrated into the mainstream of European research. For all Czech scientists, the project could help to restore self-esteem.

It is therefore important that there is no unnecessary delay in the installation of the laser in Prague. Concrete for the special housing unit can be laid only in the autumn months. If financial uncertainty forces a one-year delay to plans to lay the foundations next month, scientists who were using the facility at Garching until it closed last spring will look elsewhere. And if this happens, an important opportunity for Czech science will have been lost. □



Seismologists claim quake data being 'mis-read' as bomb test

[WASHINGTON] Publicly accessible data clearly show that a seismic event in northern Russia, which has been interpreted as a possible secret nuclear weapons test, was in fact a small earthquake, according to prominent US seismologists. But they are angry that their assessment is apparently being ignored by the White House.

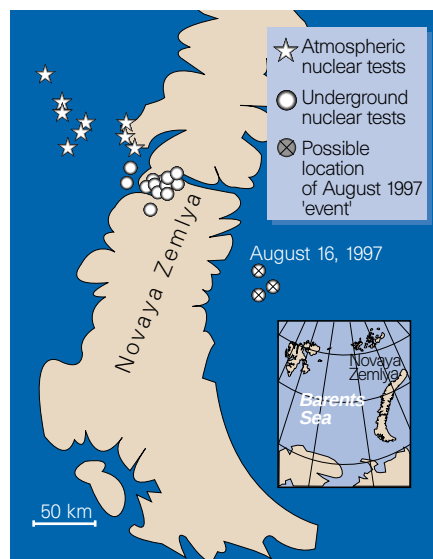
The event took place on 16 August more than a hundred miles from the Russian test site at Novaya Zemlya. The Clinton administration has refused to say whether it believes the event was an earthquake or a nuclear test.

"We're still reviewing the technical data that continue to come in on that incident," Bob Bell, a senior official of the National Security Council, said last week. "The data are not conclusive; they lend themselves to alternative explanations."

Jeff Smith, a spokesman for the Office of Science and Technology Policy (OSTP), which coordinates science policy at the White House, said that OSTP was part of an "inter-agency review" of the data. He declined to estimate when the review would be completed.

But seismologists say that data from open seismic stations in Norway, Finland and Russia already indicate clearly that the event was an earthquake, not an explosion. Not only did it take place about 130 km off the coast of Novaya Zemlya, but it also has a seismic signature that marks it as an earthquake rather than an explosion.

"I don't know of any seismologist who doesn't think that this was an earthquake," says Greg van der Vink, of the Incorporated Research Institutions for Seismology, a col-



Mapping the event: location and type of wave, measured in Norway, Finland and Russia, clearly point to an earthquake, say seismologists.

laboration of 90 universities which is supported by the National Science Foundation to operate seismology research facilities.

When seismologists study such events, they measure 'S' (or shear) waves as well as 'P' (or compressional) waves in the Earth. Underground nuclear explosions produce a signature dominated by P waves, whereas earthquakes have a significant s wave component.

According to Jeffrey Park, an associate professor of geology at Yale University, the 16 August event was characterized by shear waves. "If you were looking for a potential

nuclear explosion, this just wouldn't ring any bells," he says. Several other leading Earth scientists support Park's assessment.

But press reports suggested that the event was a nuclear explosion after an unnamed Department of Defense official told the *Washington Times*, a conservative newspaper, that it had "explosive characteristics".

The official told the newspaper that the event undermined the Comprehensive Test Ban Treaty (CTBT), because it showed either that the Russians were cheating or that the United States was technically incapable of figuring out whether they were. Supporters of the treaty hotly dispute both assertions.

President Bill Clinton announced last week that he will now take the CTBT to the Senate, where it needs to win 67 votes out of 100 for verification. Fifty-seven of the senators are Republicans, and the Republican party is officially opposed to the treaty. Observers believe, however, that Clinton has a realistic chance of winning enough support for verification.

Senators who oppose the treaty, including Jesse Helms (Republican, North Carolina), chair of the Foreign Relations Committee, and Jon Kyl (Republican, Arizona), were quick to cite the 16 August event as justification for their opposition. Supporters of the treaty are therefore urging the administration to declare that the event was not a nuclear test.

"The administration has handled the whole episode extremely badly," says Chris Paine, an analyst with the Natural Resources Defense Council, an antinuclear lobby group. "They've known for weeks" that the incident was an earthquake, he says. "But they haven't said anything".

Paine also alleges that seismologists have been slow to speak out because many depend on the Department of Defense for funding.

A group of seismologists led by Lynn Sykes of the Lamont-Doherty Earth Observatory at Columbia University, New York, abandoned plans to hold a press conference to show their evidence that the 16 August event was an earthquake. Sykes was unavailable for comment on this decision.

Jeremy Stone of the Federation of American Scientists, a group of scientists opposed to nuclear proliferation, says that the reporting of such events has followed a familiar pattern since the height of the Cold War.

"Semi-digested intelligence information is leaked from the intelligence community, whenever it seems to suggest that an alarm bell should be rung," he says. "This leads to a permanent skewing of our view on whether the other side is cheating." **Colin Macilwain**

Russian researchers brace for more cuts

[MOSCOW] A reduction in science spending of about 12 per cent is being proposed in the Russian government's draft budget for 1998, which was submitted last month to the State Duma, the lower chamber of parliament. Duma members have expressed concern about the proposed cuts.

A parallel submission to the Duma threatens to make conditions for science even more difficult by abolishing many existing laws, such as that requiring at least 4 per cent of all budget expenditures to be spent on science and at least 3 per cent on education.

"The draft budget for 1998 allocates approximately 12 per cent less for the year compared with 1997," complains Ivan Melnikov, the chairman of the Duma's committee on education and science. Under the proposals, financing of the Russian

Academy of Sciences would be reduced by 19 per cent and that of its Siberian branch by 25 per cent, while the programme for boosting research in universities would receive only half of its 1997 allocation.

News of the proposed cuts for 1998 follows a reduced budget for 1997. Melnikov says that in May the cabinet drafted a law holding back promised state expenditure in 1997, stating that education and science were to lose more than 4,000 billion roubles (US\$700 million). These cuts were rejected by the Duma, but the government nevertheless went ahead with them.

The 1997 cuts meant that general expenditures for the first half of the year were only 70 per cent of those promised, with science receiving only 58 per cent of its planned budget, says Melnikov. **Carl Levitin**

US dispute over live AIDS vaccine trials

[WASHINGTON] The leaders of a group of 50 volunteers who argue that the time is ripe for human trials of an AIDS vaccine containing live HIV claim that the US National Institutes of Health (NIH) are dragging its feet over support for such work.

But NIH leaders and other leading scientists say that, although the volunteers' goal is laudable, it is premature.

The conflict highlights growing differences between scientists and clinicians about what constitutes ethical, timely AIDS vaccine research. It has been stirred up by the high response to a journal announcement soliciting volunteers for a trial of live attenuated — or weakened — virus vaccine; 39 of those who responded are AIDS physicians.

The leaders of the group heading the volunteer effort, the International Association of Physicians in AIDS Care (IAPAC), met staff of the National Institute of Allergy and Infectious Diseases (NIAID) last week. After the meeting, IAPAC's executive director, Gordon Nary, said NIAID staff had made "very, very helpful" suggestions about what would be required for the group to develop a vaccine protocol that would be approved by the US Food and Drug Administration.

But Nary said there are still "major differences" between the volunteers and government scientists about how much animal work remains to be done before human trials of a vaccine of live attenuated virus can proceed.

"How many more deaths are going to move us to humanize science?" asked Nary, arguing that what was needed was "science in the service of humanity, as opposed to science in the service of science".

IAPAC contends that the results of research with a live attenuated monkey analogue of HIV have been sufficiently positive for live, attenuated vaccine to be tried on a small scale in humans. It plans to present a research protocol, aiming to vaccinate up to 15 volunteers by the year 2000, at a Washington conference in November.

Leading scientists and NIH officials call the group's move premature. NIAID's director, Anthony Fauci, says he admires the "altruism" of the IAPAC volunteers, but argues that they are guilty of "misrepresentation: that we have a vaccine in hand and we're ready to go; all we need is the volunteers".

He says there remain substantive obstacles to human trials, especially concern about safety. These can be alleviated only "by

doing scientific experiments in the orderly scientific fashion that they should be done".

David Baltimore, professor of molecular biology at the Massachusetts Institute of Technology and head of the AIDS vaccine research committee of the NIH, agrees. While calling the volunteers "very brave", he says: "Even if they're willing to put themselves on the line, I don't think that it's appropriate to test materials that we don't feel are reasonably certain to be safe."

Baltimore says that his concern about safety has been increased by at least two unpublished monkey studies that "raise concerns about whether there isn't a fraction of animals who receive the live attenuated virus who go on to develop disease in time".

In justifying its demand for rapid human trials, IAPAC has cited the work of Ronald Desrosiers, a virologist at the Harvard Medical School. Desrosiers has demonstrated immune protection lasting for more than seven years in monkeys vaccinated with live attenuated SIV, the monkey analogue of HIV.

Desrosiers has provided reagents and other starting materials to Therion Biologics of Cambridge, Massachusetts, which is working to develop a live attenuated vaccine. The vaccine was originally developed by Desrosiers and exclusively licensed to Therion.

But Baltimore argues that more urgent than human trials is an understanding of the basic biology of live attenuated virus, without which it is nearly impossible to assess the appropriateness of human trials. He says the NIH's AIDS Vaccine Research Advisory Committee is emphasizing support for basic research in this area, for instance through backing such work in 'innovation grants' announced on Monday (29 September).

The fact that live attenuated virus has produced the best results in monkeys means that understanding how it acts needs to be "a very high priority" in government funded research, Baltimore says.

IAPAC solicited volunteers through the August issue of its journal, which goes to 5,500 members in 42 countries. Charles Farthing of IAPAC, medical director of the AIDS Healthcare Foundation in Los Angeles, calls further primate trials "unnecessary, expensive and time-consuming".

Concerns about live attenuated HIV vaccine include the possibility that it might revert to fully virulent form in recipients, and that low-level persistence of the virus might cause AIDS decades after vaccination.

There are also concerns that DNA from the injected virus, which incorporates itself in the host genome, might induce other diseases, for instance by activating cancer-causing genes, and that infants might be harmed by low-level infection induced by vaccination in their mothers.

Meredith Wadman

German minister tipped to head UN body

[LONDON] Klaus Töpfer, Germany's minister for housing and a former environment minister, is being tipped as a leading candidate to succeed Elizabeth Dowdeswell as head of the United Nations Environment Programme (UNEP).

Töpfer's name has informally attracted broad international support, particularly from the United States, Scandinavia and Germany, which is keen for one of its nationals to head a UN agency. "Töpfer's got a lot more status and experience than the others in the field," says one source at an environment ministry in a European Union member state.

Other names that have been mentioned include Michael Cutajar, head of the UN Climate Convention, and John Gummer, Britain's former environment secretary. But none of the three is likely to be supported by Canada, which sponsors Dowdeswell.

Dowdeswell's contract expires at the end of this year, and a consensus is emerging among member countries that UNEP needs a new face. "There are a lot of fundamental problems with UNEP. It needs someone with fresh ideas and vision who can invigorate the place," the source says.

UNEP's controversial 'turf wars' with its constituent conventions have also damaged its image. Some conventions — such as the Biodiversity Convention (see *Nature* 389, 5;



Töpfer: 'has much more status and experience than other candidates'.

1997) and the Convention on Trade in Endangered Species (CITES) — favour greater autonomy for themselves, including the right to deal directly with member countries. But UNEP opposes this approach.

Some countries wanted Dowdeswell replaced last year. But she was granted a one-year extension by the then UN secretary general, Boutros Boutros Ghali, partly because no candidate had been nominated, and partly because Boutros Ghali was himself about to leave his post, and so was unable to devote time to selecting a candidate.

One reason for a relative lack of enthusiasm for the UNEP post is that it is located in Nairobi. "Had the job been in Geneva, there would have been no shortage of candidates," says one source.

Under United Nations rules, the secretary general takes soundings from member countries before proposing a single candidate who he anticipates will attract widespread support. This name then goes forward for ratification before the general assembly in New York.

Ehsan Masood

British guidelines set out standards for genetic tests

[LONDON] The British government has published a set of voluntary guidelines setting out the conditions — including commitments to quality, confidentiality and the counselling of customers — that it expects companies selling genetic tests directly to the public to comply with.

The guidelines have been drawn up by the Department of Health's Advisory Committee on Genetic Testing. The committee's chairman, Marcus Pembrey, professor of paediatric genetics at the Institute of Child Health in London, said last week that a voluntary code was considered preferable to legally enforceable regulations.

Two British companies, University Diagnostics Limited and Leeds Anti-natal Screening Services, offer mail-order tests for cystic fibrosis mutations to potential parents, with all other forms of genetic testing taking place through the National Health Service. But the number of companies involved is expected to increase in the next few years, as is the range of tests on offer.

Tessa Jowell, the public health minister, said last week that the government hoped that all such companies would agree to follow the guidelines as this would ensure that "such services are available without need for more formal controls". One requirement is that all testing laboratories would be accredited by a recognized body. □

'Policy vacuum' worry as Australian minister quits

[ADELAIDE] Australia's Coalition government, led by prime minister John Howard of the Liberal party, has been plunged into crisis following the forced resignation last Friday (26 September) of Peter McGauran, Minister for Science and Technology, in the wake of the departure of two other ministers.

McGauran first denied charges of falsely claiming personal allowances for official travel but then admitted some, including having billed the government for a charter flight on which National party officials — not himself — had travelled. McGauran became the third ministerial victim of the controversy over travel irregularities following Opposition attacks, and the first science minister ever to resign.

The government's disarray over the resignations means that the replacements for McGauran and the other ministers are not expected to be announced for a week.

McGauran had served 18 months in the post, and overseen two rounds of cuts to research funding, but was respected by leaders of the scientific community. Although science policy was not a cause of the crisis, McGauran's actions have made it one of the focal points of the political controversy.

Only a month ago, he announced a new research reactor costing A\$300 million (US\$215 million), the largest government commitment to a science and technology



McGauran: oversaw two rounds of cuts.

facility in Australian history (see *Nature* 389, 109; 1997).

According to John Stocker, the chief scientist, speaking before McGauran's resignation, the minister had not sought advice on the reactor proposal from him or the official advisory body, the Australian Science and Technology Council, which Stocker chairs and which carried out two substantial studies of nuclear fuel, science and technology 12 years ago.

Martyn Evans, the Opposition spokesman for science, says the Labor party and the minority Democrats are seeking a Senate inquiry into the reactor decision. Speaking at the congress of the Australian and New Zealand Association for the Advancement of Science, being held in Adelaide this week, Evans said: "Further delay and confusion while a new minister settles in will add to the pressure on the research community."

Scientists are alarmed about the vacuum created by the departure of their minister at a time when contentious recommendations of major reports were due to be resolved by McGauran and his senior minister, John Moore (see *Nature* 388, 509 & 819; 1997).

Peter Pockley

Nobel laureates face libel suits from 'water memory' researcher

[PARIS] The long-running saga of research on the 'memory of water' has reopened with a splash, with libel suits being filed against three scientists — including two Nobel prizewinners — by Jacques Benveniste, the French researcher who claimed in 1988 to have shown that extreme dilutions of antibody solutions could retain their biological activity (see *Nature* 333, 816; 1988).

The charges are based on statements made by the scientists in January in the newspaper *Le Monde* which suggested that Benveniste's research may have been fraudulent.

A court battle is now on the cards. This week, lawyers representing two of those being sued — Georges Charpak, who won the Nobel prize for physics in 1992, and his colleague Claude Hennion, from the School of Industrial Physics and Chemistry in Paris — said they intend to fight the libel charges, first on procedural grounds, but if necessary by investigating Benveniste's research.

A spokeswoman for a Paris-based law firm, Kahen and Associates, says that, as a

civil servant, Benveniste should have filed a penal suit and not a civil one. If this is confirmed by the court, she adds, it would annul the procedure, and prevent Benveniste from suing again on the basis of the *Le Monde* articles — although he could bring new charges on any statements made by the scientists elsewhere.

But she adds that, if this first approach fails, the law firm is ready to counterattack in other ways. It could argue that Charpak and Hennion made their statements "in good faith", or seek to prove that Benveniste did indeed commit fraud.

Lawyers representing François Jacob, who shared the 1965 Nobel prize for physiology or medicine and is also being sued, were unavailable for comment.

Benveniste describes the attempt to halt the suits on legal grounds as "pathetic", and adds: "It is incredible that a Nobel prizewinner, with the sense of responsibility that this [status] carries, could affirm that his scientific colleague was a fraudster and then try to get off with legal arguments."

Benveniste says that none of the scientists has provided proof of fraud, and he decided to sue to defend his honour and professional integrity. He claims that he wrote to the scientists earlier this year saying that if they retracted the statements he would not take further action, but that he received no reply. He says he will seek damages of FF100,000 (US\$17,000).

The controversy includes Benveniste's more recent research, in which he claims to be able routinely to transmit biological activities to water or cultured cells electronically, to store such signals on computer discs, and to send them over the Internet.

Benveniste, whose laboratory was closed in 1994 by INSERM, the national biomedical research organization, now operates from the privately funded Digital Biology Laboratory at Clamart near Paris. He admits difficulty in raising the laboratory's running costs of FF100,000 a month, but predicts that "when it takes off it will be the next Microsoft". Declan Butler

NIH inquiry fails to find culprit of contamination

[WASHINGTON] The US National Institutes of Health (NIH) are not to be stripped of their licence to use nuclear materials, despite the efforts of a pregnant scientist who was contaminated with radioactive phosphorus while working there (see *Nature* 377, 568; 1995).

The Nuclear Regulatory Commission (NRC) has denied a petition from Maryann Wenli Ma and her husband, Bill Wenling Zheng, that NIH's licence to use nuclear materials be revoked or suspended.

After the discovery of Ma's exposure in June 1995, it was found that 26 others — including Zheng — had been exposed to radiation from a contaminated water cooler at NIH (but not through the consumption of Chinese food that had been left in a refrigerator, as originally thought).

Investigators from the NRC, working jointly with the Federal Bureau of Investigation and the NIH police, concluded that the contamination was "deliberate", but could not identify the perpetrator.

The NRC said in a statement that, normally, "radiation overexposures of this sort would be subject to significant enforcement action". But it added that it would exercise discretion in this case because there was no evidence that NIH had contributed directly or indirectly to the deliberate misuse of licensed material.

Also, NIH "could not reasonably have foreseen" malicious misuse of materials by an employee, and had cooperated fully in the investigation.

The NRC did say, however, that NIH had violated several requirements for the security and control of radioactive materials, but that the agency had since made "significant efforts to improve its control".

David Marshall, a lawyer representing Ma and Zheng, says that the decision "allows the [NIH] to continue to put its employees and the public at risk with impunity". The lawyers are considering filing a civil suit for damages in the federal court. Although Ma's son David was born apparently healthy, "Dr Ma and her child will live for years with the dangerous and uncertain effects of radiation poisoning", says Marshall.

The NIH has issued a statement welcoming the NRC conclusion. "It is particularly significant," it said, that the NRC found the contamination "was not the result of faulty compliance with security requirements for radioactive materials".

But the statement also pointed out that NIH had nevertheless "consistently tightened its standards for the security and use of such materials".

Meredith Wadman

UK nuclear waste company gets extended lease of life

[LONDON] Britain's embattled nuclear waste disposal company Nirex has been given a one-year lease of life under a new part-time chairman, quashing speculation that it was on the verge of being wound up or merged with one of its shareholder companies.

A meeting of the Nirex board last week agreed to keep the company intact with no further job losses for the next 12 months and to maintain its £30 million (US\$49 million) annual research budget. A final decision on Nirex's future will now depend on the new government's yet-to-be-announced plans for disposing of nearly 300,000 cubic metres of radioactive waste from nuclear power plants.

Speculation had been mounting before the meeting that Nirex was to merge with its largest shareholder, British Nuclear Fuels (BNFL), which owns 40 per cent of Nirex.

But the board chose to bring the two closer together by appointing David Bonser, BNFL's director of waste management and decommissioning, as Nirex's new chairman.

The company's future was put in doubt last April (see *Nature* 386, 423–424; 1997) after the previous government blocked its plans to dispose of nuclear waste at a proposed underground repository at Sellafield in north-west England. Nirex had spent more than £200 million investigating the site. The choice of Sellafield had been opposed by the local authority and environmentalist groups, which also agree that waste must be retrievable.

With no immediate prospect of an alternative site to investigate, Nirex had little

choice but to begin a rapid programme of retrenchment, and shed more than half of its 200 staff, including many scientists. Some believed that cuts would continue. But in a statement, Bonser said that he looked forward to "building for the future on [Nirex's] scientific excellence".

Environmental groups have reacted coolly to the appointment, arguing that it suggests a 'business as usual' approach in an agency that has come under fire for its secrecy and for the perceived lack of public involvement in its decision to choose Sellafield. Helen Wallace, a senior scientist with Greenpeace, says Bonser's appointment will only reinforce such an impression. "Nirex is not going to gain any public support by appointing a nuclear industry insider."

But one former government adviser on radioactive waste policy, while accepting that the choice of chairman "is not a step that will create public confidence", says that the one-year moratorium on further changes and the appointment of a part-time chairman suggests that the nuclear industry wants to keep a watching brief over Nirex while giving the new government flexibility in planning for the future.

If it chose to do so, the government could still retain Nirex as the agency responsible for radioactive waste disposal, he says. But the board's decision does not foreclose the other option — preferred by environmentalist groups — of setting up a new, independent and publicly accountable agency.

Ehsan Masood

'Government should decide on disposal site'

[LONDON] Nirex, Britain's nuclear waste disposal agency, has made clear that it does not want responsibility for choosing a new waste-disposal site — even if the government decides that it should retain its independence.

According to John Holmes, Nirex's director of technical services, the government and not Nirex should take the lead in any future site-selection process; and, once a decision is made, it should be protected from politically motivated interference.

Nirex took much of the blame for the choice of the

Sellafield site, even though the decision is believed to have been fully endorsed by the late Nicholas Ridley, environment secretary in the then Conservative government, only to be overturned by his successor, John Gummer.

Holmes was speaking last month at a discussion organized by the Geological Society of London on the future of radioactive waste disposal in Britain. He acknowledged that Nirex had made mistakes, by, for example, not discussing its plans with more scientists outside the company.

Despite planning for a

sealed underground repository, one Nirex official has revealed that Nirex's eventual choice of repository would have had a degree of retrievability almost by default.

The safest, but most expensive and least retrievable, option would have been to build a deep repository that would be filled with concrete every time waste was placed at weekly or monthly intervals. A cheaper, more retrievable method would have been to fill a repository with concrete only when it was full of waste. "Nirex was tending towards this option," he says. **E.M.**

'No net cost in cutting carbon emissions'

[WASHINGTON] The United States could cut carbon emissions to 1990 levels by 2010 with no net cost to the nation's economy, according to a Department of Energy (DOE) report. This reduction could be achieved by increasing the country's use of energy-efficient technology, says the report, which was released in Washington last week.

But the federal government would first have to invest "far beyond current efforts" in developing and promoting that technology. In fact, say the study authors, meeting the goal will require a "major effort to reduce carbon emissions through federal policies and strengthened state programmes, and very active private sector involvement".

The report, the product of a year-long study by five DOE laboratories, looked at ways for some 200 technologies to improve energy efficiency in four sectors of the economy: buildings, transportation, industry and electric utilities.

It concluded that actions such as switching to natural gas instead of electricity for residential use, replacing coal-burning plants with plants that use natural gas, and increasing reliance on biofuels such as ethanol for transportation, could reduce the nation's energy bill by between \$50 bil-

lion and \$90 billion a year. That sum would be enough to offset the cost of developing and switching to the new technologies, according to the report.

The study's analysis relies heavily on the assumption that industry would also be allowed to trade carbon emissions on the open market. Assuming a trading price of \$50 per tonne of carbon and greater use of energy-efficient and low-carbon technology, emissions could be reduced by 390 million tonnes per year, bringing them back to 1990 levels. This strategy, combined with next-generation energy technologies, could lead to even greater reductions in the next quarter-century, according to the DOE analysts.

Their conclusions are in line with a 1992 study by the National Academy of Sciences and a 1991 assessment by the congressional Office of Technology Assessment.

The DOE report, which was reviewed by outside scientists and economists, gives ammunition to those hoping that the United States will commit itself to mandatory greenhouse gas reductions at the international conference in Kyoto, Japan, in December.

In Kyoto, the United States will be under political pressure from European and other countries to agree to reduce greenhouse



Solar futures? Greater efficiency could cut energy bills, offsetting the costs of new technologies.

emissions to at least 1990 levels by 2010 (see *Nature* **388**, 614; 1997).

But the Clinton administration is still playing its cards close to its chest as it prepares for the meeting, and has not disclosed any proposed targets or timetables.

Industry lobbyists have fiercely opposed mandatory targets, and the new DOE report is unlikely to sway them. These critics have claimed that reducing greenhouse emissions to 1990 levels would cost Americans as much as \$2,000 per household in 2010. The American Petroleum Institute (API) estimates that capping carbon-dioxide emissions in the short term could reduce US gross domestic product by two to four per cent — the equivalent of \$200 billion to \$350 billion per year.

William O'Keefe, API's executive vice-president, who also chairs the industry-sponsored Global Climate Coalition, last week dismissed what he called "ongoing administration efforts to perpetuate the self-contradictory myth that 'free lunch' technologies are readily available" to solve the greenhouse emissions problem.

While the Global Climate Coalition supports government investment in new energy technologies, said O'Keefe, "neither the price nor the timing of major technological breakthroughs can be predicted". He also indirectly accused the DOE of self-interest in arguing for increased investment in clean technology. "The DOE national laboratories may need a new mission now that the Cold War has ended," he said.

Meanwhile, government officials, scientists and corporate leaders will meet next week to discuss global warming policy at Georgetown University in Washington. President Bill Clinton, who is hosting the conference, may use the occasion to be more specific about US plans for meeting greenhouse targets.

Alternatively, Clinton may leave the specifics of the plans to the US representatives attending a pre-Kyoto meeting in Bonn, Germany, later this month. **Tony Reichhardt**

Japan's ministries argue over greenhouse target

[TOKYO] Japan's efforts to set a binding target for a cut in greenhouse gas emissions were set back last week when the three government ministries involved failed to agree on a joint position.

Prime Minister Ryutaro Hashimoto had asked the three ministries — the Ministry of International Trade and Industry (MITI), the Environment Agency and the Foreign Ministry — to agree on a target by the end of September.

But agency officials are still locked in negotiations (see *Nature* **387**, 641; 1997), and it is widely believed they are unlikely to reach an agreement until the middle of October at the earliest.

Hashimoto wants Japan to take a strong stance at the review conference of the UN Framework Convention on Climate Change, which takes place in Kyoto in December.

The Japanese govern-

ment has already said it considers a European Union proposal for a 15 per cent cut from 1990 emissions levels by the year 2010 to be unrealistic.

In an interim report submitted to the prime minister last week, the Environment Agency proposed that emissions of carbon dioxide and other greenhouse gases be cut by 7 per cent from the 1990 levels by 2010, and the Foreign Ministry suggested a 6.5 per cent cut. But MITI is said to want no reduction from 1990 levels.

At a joint council of a government advisory panel last week, MITI submitted a plan to cut carbon-dioxide emissions to 1990 levels by 2010 through doubling nuclear energy production and using alternative energy sources such as solar power.

MITI officials say further reductions in carbon-dioxide

emissions beyond that level could seriously damage the economy, and an overall target as proposed by the Environment Agency would create major difficulties.

MITI's position is based on the fact that the measures required to reduce greenhouse gas emissions — such as reducing traffic in Tokyo — would require the complex revision of regulations over and above those concerning energy conservation.

The ministry has also clashed with the Environment Agency about whether to set a per capita reduction of greenhouse gases, favoured by MITI, or the flat-rate reduction that the Environment Agency wants.

With only two months left until the Kyoto meeting, which Japan is chairing, the delay is expected to attract criticism from other signatory nations which have been calling on Japan to clarify its reduction target. **Asako Saegusa**

Czech science fights against flood costs

[PRAGUE] Czech scientists are waiting nervously to learn what sacrifices they will be required to make to help to pay for the central European floods which covered a third of the Czech Republic in July.

The government's budget proposal is due to be presented to parliament this week. It is likely to propose cuts in funding for research which, say scientists, will make it difficult for them to bring to fruition ambitious plans to bring Czech science up to European Union (EU) standards.

Researchers are particularly worried that plans to establish what could be the Czech Republic's first international large-scale facility—a high-power iodide laser, of a type known as Asterix—could be disrupted.

Karel Jungwirth, a plasma physicist who helped to negotiate the laser's transfer from the Max Planck Institute for Quantum Optics in Germany, and vice-president of the Czech Academy of Sciences, has won promises of support from grant agencies and from special government funds (see box).

But he fears that next year's science budget may throw the timetable into disarray. "It needs to be up and running quickly so we don't lose the circle of users," he says. "But more significantly, it will be an important signal to the young that science has a future in the Czech Republic."

This is a major issue for the Czech scientific establishment, as well as the government. The sudden arrival of capitalism has seen the brightest young scientists and students turn their backs on research careers, and the average age of researchers in the Czech Republic is rising alarmingly.

Wages are particularly unattractive to young researchers. The average academic salary of 10,000 Koruna (US\$315) per month is significantly lower than the average commercial salary, and all university contracts, even those of professors, are temporary.

Scientists are campaigning hard for the government to stick to a 1996 election pledge to give science top priority after defence. But



Troubling waters: science was hit directly as well as indirectly by the flooding in July. The Czech Academy of Sciences' Institute of Experimental Botany (pictured) in Olomouc, east of Prague, was under 2.5 metres of water for more than two hours. All its equipment, most of which was bought in the last two years, was destroyed.

the government, which is trying to steer the country towards membership of the North Atlantic Treaty Organization and the EU, has already shown itself unable to provide financial backing for its good intentions.

In April, the government approved its 'Principles in the field of research and development, 1997'. This promises a re-organization of research to bring it into line with EU norms. The principles pledge an increase in total state investment in research from 0.4 to 0.7 per cent of gross national product by the year of entry into the EU, which it hopes to join within seven or eight years. There are also promises to improve cooperation between academic research and industry, and to increase investment in industrial research.

Only in the past two years has the Czech Republic begun to invest significant funds in applied research, but these still total less than 15 per cent of the total state research budget.

The principles promise to continue development of initiatives already under way. These include more programmes for targeted research, promotion of university research which had been largely confined to the Academy of Sciences research institutes during the 40 years of Communism, and



transfer of more institutional funding to grant agencies.

But, within two weeks of approving these principles, the government was forced by an economic collapse following a run on the Czech currency to introduce an emergency budget. This was followed a few weeks later by a second emergency budget.

As a consequence, the Academy of Sciences saw its budget, which had this year received its first real-terms increase since the fall of Communism, suddenly slashed by more than 10 per cent. Universities and grant agencies suffered similar cuts. Potentially most damaging was the government's decision to freeze salaries for all public appointments, including academics.

The academy feels that it has already trimmed all its fat, having closed 22 research institutes and shed more than half its staff following the halving of its budget in 1993 (see *Nature* 368, 386; 1994). Teams of experts, nearly half of them foreign, have completed detailed evaluations of the academy's remaining 59 institutes, and their recommendations are being incorporated into the institutes' plans. "This system is unique in eastern Europe," says Ladislav Pivec, secretary of the academy's appraisal committee.

There could be worse cuts to come. "The floods came," says Jungwirth despairingly, "and we don't know how we will cope with what this could do to our budget."

Josef Syka, vice-chairman of the government's Research and Development Council, suggests that the government will propose to maintain the total science budget at Kc7 billion, the level the current budget reached after the two emergency budgets cut it by nearly 15 per cent.

Alison Abbott

Academy focuses on bringing laser into use

[PRAGUE] The iodide laser was sold to the Czech Academy of Sciences last spring by the Max Planck Institute for Quantum Optics in Garching, near Munich, for a token cost of DM1 (US\$0.56). It was one of a series of Asterix lasers developed in Garching and optimized for laser and X-ray plasma experiments. It had additional support from the European Commission, and

was used by scientists from many countries.

If financing is maintained, the laser is to be housed in an academy campus on the outskirts of Prague and run by the academy's institutes of physics and plasma physics, which are also on the campus. The Max Planck Institute decided to get rid of the laser after reorienting its research plans. Under its

contract with Prague, its scientists will continue to have free access to the laser when in operation.

The Czech academy hopes to maintain the laser's status as an international facility through domestic and foreign support. But it knows that if there is too long a delay in getting the laser operational, its circle of users will be lost.

A.A.

Irreproducible pursues the improbable

[BOSTON] This year's Ig Nobel Prize ceremony — to be held next week at Harvard University under the appropriate title of 'The Big Bang' — seems likely to generate more pyrotechnics than usual, with rights to the 'Ig' trademark itself being hotly contested.

Marc Abrahams, editor of the *Annals of Improbable Research* (AIR), has organized the event for the past seven years. But George Scherr, a microbiologist and publisher of a rival science humour magazine, the *Journal of Irreproducible Results* (JIR), has sought to gain control of the term 'Ig Nobel Prize' in an application to the US Patent and Trademark Office.

Scherr is also suing Abrahams, a former editor of JIR, for \$4.2 million, accusing him of unfair business practices, trademark infringement, conspiracy to defraud, racketeering and other offences. The world of irreproducible results has declared war on improbable research, leaving the future of dubious science hanging in the balance.

In August, the trademark office rejected Scherr's ownership claim for the term 'Ig Nobel Prize', but Scherr has six months to appeal against that decision. He declines to discuss the trademark issue or his grievances with Abrahams other than to say: "Nature is not the proper venue for this discussion. Legal disputes should be resolved in a court of law."

Abrahams is more forthcoming, maintaining that Scherr has no legitimate

Your Honour, this is a highly serious dispute about who owns the copyright of the custard pie.



claim to the 'Igs'. Alexander Kohn — an Israeli physicist who co-founded JIR in 1955 and then co-founded AIR (with Abrahams) in 1994, shortly before his death — presented several 'Ignobel' prizes in a 1968 JIR article that contained perhaps the first published reference to the concept. Kohn encouraged Abrahams to award such prizes at the new ceremony he started in 1991.

In *The Scientist* in 1996, Scherr argued that "neither Abrahams nor AIR were ever associated with the Ig Nobel Prize". But this has been challenged by Nobel laureates Richard Roberts and Dudley Herschbach, both of whom thought they had participated in numerous Ig ceremonies hosted by Abrahams. "Could it be that Scherr and I inhabit parallel universes?" Roberts asked.

Abrahams is equally baffled by the lawsuit, which seeks to prevent him from publishing any journal that (at least in Scherr's mind) is likely to be confused with JIR. Abrahams contends that the names AIR and JIR are not, as Scherr argues, "confusingly similar".

But to further dispel any possible misunderstanding, Abrahams has printed a disclaimer in AIR stating that the magazine "is in no way associated with that publication [JIR] or with its publisher".

Abrahams says he cannot understand why he has been charged with "conspiracy" as no one else is mentioned in the lawsuit. "Evidently I've conspired with myself," he says. He is also astounded by the huge sum that Scherr hopes to collect, given that AIR has only 2,000 subscribers who pay just \$23 annually. "Why \$4.2 million? Why so low?" Abrahams jokes. "Doesn't Scherr respect me?"

On a more serious note, he has enlisted the help of Nobel laureates Herschbach, Roberts and William Lipscomb to launch the "Strategic AIR Defense Fund" to support his legal fight.

Abrahams speculates that his science humour colleague is making a serious bid for an Ig Nobel Prize. The situation is not one Abrahams imagined three years ago when he started his humour magazine. "I just want to be able to write and edit some funny things about science and maybe do some good along the way," he says. "I have no idea what Scherr wants."

Steve Nadis

Netanyahu bolsters opposition to dismemberment of ministry

[JERUSALEM] Israel's Prime Minister, Binyamin Netanyahu, has made it known that he opposes plans to dismantle the country's Ministry of Science, following an outcry from the scientific community.

The proposal to break up the ministry and transfer most of its grant-making authority to a national research and development council was made to the government last month by the minister of science, Michael Eitan. Eitan argued that such a reorganization would eliminate unnecessary administrative costs, and increase the amount of public funds available for basic and long-term research.

Eitan has subsequently admitted to the science committee of the Knesset — Israel's parliament — that he might have made a tactical error in proposing the dismemberment of the ministry. He has therefore offered to withdraw that part of his plan. But he continues to advocate the establishment of a single council to coordinate the government's science policy.

This body would coordinate all publicly

funded support for research and development, including both the directed long-term research supported by the current ministry and the applied research funded by the Ministry of Industry and Commerce and other ministries.

Eitan suggested that administrative work associated with grants should be carried out by the industry ministry, and that a minister of science should continue to bear responsibility for government science policy, but working from the prime minister's office rather than heading his or her own ministry.

Eitan claimed that the finance ministry had promised that, if the reorganization programme was adopted, his ministry's entire budget — some US\$220 million — would be made available for basic and long-term research grants. More than half of the science ministry's budget is at present spent on administration and other research and education programmes.

But his proposals generated a protest from most of Israel's scientists. A petition opposing the plan collected the signatures of

more than 700 scientists at all seven Israeli research universities. Three former ministers of science voiced objections, as did the deputy minister of defence, Silvan Shalom, who is due to take over the science ministry next summer, under a rotation arrangement with Eitan (see *Nature* 388, 221; 1997).

"Decentralizing authority over science can harm science, especially in the long run," argues Professor Arie Admon of the Technion, the Israel Institute of Technology. Without an interested party such as a government ministry to fight for the science budget, he says, funding will inevitably decrease in the long run. Admon also claims that the science ministry has been highly successful in planning and coordinating research among Israel's universities.

Netanyahu's opposition to the dismemberment of the science ministry was announced by his science adviser, Israel Hanukoglu. Criticism has also come from the Knesset's science committee and from the former minister of science Ze'ev Binyamin Begin.

Haim Watzman

CERN council fails to agree on new director-general

[MUNICH] Last week's meeting of the committee of council of CERN, the European Laboratory for Particle Physics in Geneva, Switzerland, failed to come up with a new chief. The committee was aiming to choose a successor to Christopher Llewellyn Smith, whose term of office as director-general expires at the end of this year.

The three official candidates being considered are Albrecht Wagner, director of research at DESY, Germany's national research centre for particle physics; Robert Aymar, director general of ITER, the next-generation fusion reactor; and Walter Hoogland, a former director of research at CERN who is now at the Nikhef Institute of Particle Physics in Amsterdam.

Another attempt to select a new director-general will be made at a special committee of council meeting in November.

Cray wins trade ruling on supercomputer deal

[WASHINGTON] The US International Trade Commission ruled last week that US-based Cray Research was financially injured by Japanese rivals which won a bid last year to sell a supercomputer to the National Center for Atmospheric Research (NCAR) (see *Nature* 382, 385; 1996).

Although the Japanese-based NEC Corporation won the bid, subsequent complaints by Cray and US politicians led the National Science Foundation, which funds NCAR, to scrap the deal.

In March, the US Commerce Department found that both NEC and Fujitsu were guilty of attempted price "dumping" in the NCAR bid. That decision, coupled with last week's ruling, clears the way for the United States to charge stiff duties on future sales of vector supercomputers by the two companies. In the case of NEC, these duties could amount to 454 per cent of the price of a system. The company has said that it intends to appeal against the decision.

India celebrates rocket launch

[NEW DELHI] India last week confirmed its presence in the global space business with the third consecutive successful flight of the Polar Satellite Launch Vehicle, its most powerful rocket. The launch is likely to raise India's standing in the eyes of South Korea, India's first foreign client in the space business; a 110-kg satellite built by South Korea is due to share a ride on the launch vehicle's next flight in 1998.

In the latest flight, the 294-tonne rocket

blasted off from the Sriharikota launch pad on the east coast and placed in an 817-km polar orbit the 1,200 kg remote-sensing satellite IRS-1D — the heaviest spacecraft launched by a domestically made rocket from Indian soil. The prime minister, Inder Kumar Gujral, who witnessed the launch, described it as his "biggest moment" in the golden jubilee year of India's independence. He congratulated the Indian Space Research Organization for having achieved self-reliance in launch capability. The three earlier satellites in the IRS series were launched by Russian rockets.

German research forum was 'complete fiasco'

[MUNICH] A scientific exhibition held in Leipzig last week and billed by the federal research ministry as "the world's biggest research forum" attracted such a low attendance that it was called a "complete fiasco" by the German press.

The *Forschungsforum '97* exhibition — organized by the research ministry which paid a total of DM2.5 million (US\$1.41 million) towards its installation — was intended to bridge the gap between researchers, industry and the public.

But one of the 2,000 or so researchers representing research organizations, universities and private industry said the exhibition was "a complete waste of time and money". And although Germany's national research centres together presented 72 exhibits, an average of only two inquiries was registered at each of them during the week.

"We were getting so bored behind our deserted stands, we soon began to visit each other instead of talking to visitors," said a participant from one of the national centres. Some researchers closed their stands and left the exhibition early.

Rice research director quits after two years

[LONDON] The director general of the International Rice Research Institute (IRRI) in Los Baños, Philippines, has resigned midway into his five-year contract. George Rothschild, who is credited with encouraging more Asian countries to fund IRRI's work, says he is leaving on 1 December for "personal reasons related to retirement from professional life".

Rothschild, an entomologist from Australia, has presided over one of the most difficult periods in the history of the 16-member Consultative Group on International Agriculture Research, of which IRRI is the largest member. At the start of the year, 'donor fatigue' from Western countries led to redundancy for half of IRRI's 1,000 staff, and a projected US\$6 million cut in its annual US\$40 million budget.

Fund will aid research into insurance losses

[LONDON] The UK government and Britain's insurance industry have launched a fund to stimulate research proposals to help the industry sharpen up its calculations of losses from natural hazards. The TSUNAMI initiative is worth £960,000 (US\$1.5 million) and is being supported equally by the Department of Trade and Industry and an eight-strong insurance industry consortium including Commercial Union, and the Royal and Sun Alliance Group.

The initiative is being led by Dougal Goodman, acting director of the British Antarctic Survey in Cambridge. Goodman says one of the aims is to help actuaries incorporate current research in their short-term forecasting of extreme climate events.

German fraud case professor dismissed

[MUNICH] Marion Brach, the molecular biologist who admitted in May having systematically fabricated data in research papers in Germany's biggest scientific fraud case (see *Nature* 387, 442; 1997), has been dismissed from her post as professor at the University of Lübeck. She had been suspended from her position since June, while enquiries were taking place. Her former colleague and partner, Friedhelm Herrmann, is also being charged in the same fraud case, but he denies any wrongdoing (see *Nature* 389, 105; 1997).

France increases 1998 research spending

[PARIS] The French government will increase spending on research next year to FFfr53 billion (US\$8.9 billion), it said last week.

At first glance, this sum represents a rise of only 1.4 per cent on the current year — roughly in line with the anticipated rate of inflation. But more than FFfr2 billion of this year's budget was held over from 1996, and so next year's budget represents a real increase of 6.2 per cent over the planned 1997 budget.

The precise distribution of next year's total between different research activities was due to be announced this week by Claude Allègre, the minister of research and education. Increasing the number of scientific jobs is expected to be a priority, both in large research bodies such as the Centre Nationale de la Recherche Scientifique and in universities.

Other beneficiaries are likely to include programmes to support innovation in industry. Some of the additional money needed for such goals is likely to be found by reducing the funding available for investment in large-scale scientific equipment (see *Nature* 388, 7; 1997).

Cloning and bioethical thinking

Sir—Axel Kahn reminds us, rightly, that Kant's famous principle states: "Respect for human dignity requires that an individual is never used... exclusively as a means", and suggests that I have ignored the crucial use of the term 'exclusively' (*Nature* 388, 320; 1997).

I did not, of course, and I am happy with Kahn's reformulation of the principle. It is not that Kant's principle does not have powerful intuitive force but that it is so vague and open to selective interpretation, and its scope for application is consequently so limited, that its utility as one of the "fundamental principles of modern bioethical thought", as Kahn describes it, is virtually zero.

Kahn himself rightly points out that debates about the moral status of the human embryo are debates about whether embryos fall within the *scope* of Kant's or any other moral principles concerning persons; so the principle itself is not illuminating in this context. Applied to the creation of individuals which are, or will become, autonomous, it has limited application. True, it rules out slavery, but so do other principles based on autonomy and rights.

If you are interested in the ethics of creating people, then, so long as existence is in the created individual's own best interests, and so long as the individual will have the capacity for autonomy like any other, the motives for which the individual was created are either morally irrelevant or subordinate to other moral considerations.

So even where, for example, a child is engendered exclusively to provide 'a son and heir' (as in many cultures), it is unclear how or whether Kant's principle applies. Either other motives are also attributed to the parent to square parental purposes with Kant, or the child's eventual autonomy, and its clear and substantial interest in or benefit from existence, take precedence over the comparatively trivial issue of parental motives. Either way, the "fundamental principle of modern bioethical thought" is unhelpful.

I am therefore at a loss to know why Kahn invokes it with such dramatic assurance or how he thinks it applies to the ethics of human cloning. It comes down to this: either the ethics of human cloning turn on the creation or use of human embryos, in which case, as Kahn himself says, "in reality the debate is about the status of the

human embryo" and Kant's principle must wait upon the outcome of that debate; or it is about the ethics of producing clones that will become autonomous human persons.

In the latter case, as David Shapiro rightly comments (*Nature* 388, 511; 1997), the ethics of their creation are, from a Kantian perspective, not dissimilar to that of other forms of assisted reproduction or indeed to the ethics of the conduct of parents concerned exclusively with producing an heir or preserving their genes or, as is sometimes alleged, making themselves eligible for public housing — and debates about whether these are *exclusive* intentions are sterile or irresolvable.

When Kahn asks: "Is Harris announcing the emergence of a revisionist tendency in bioethical thinking?", the answer must be rather that I am pleading for the emergence of "bioethical *thinking*" as opposed to empty rhetoric of invoking resonant principles with no conceivable or coherent application to the problem at hand.

John Harris

*Institute of Medicine, Law and Bioethics,
University of Manchester,
Oxford Road, Manchester M13 9PL, UK*

European centres for disease control

Sir—The US Centers for Disease Control (CDC), in Atlanta, Georgia, are a federal organization with two missions: disease control and research. Its size (several thousand people) and centralized concept make it a powerful tool for addressing the challenge of emerging and re-emerging infectious diseases (ERID), through one of its main departments, the National Center for Control of Infectious Diseases.

There are no comparable structures to the CDC in Europe. Nevertheless, such a centralized centre could boost and coordinate European efforts to research and control these diseases. Although networks can help, they are no substitute for a real centre, with researchers from different countries interacting daily. The European Laboratory for Particle Physics (CERN) and the European Molecular Biology Organization (EMBO) are successful examples. The proposed centre, whose title might be European Centre for Control of Infectious Diseases (ECCID), should have a similar international status.

Apart from control and epidemiological surveillance, ECCID should have the following components. First, basic research,

because medicine no longer controls the ERID problem. Second, training. Third, strong connections with developing countries, in which the ERID problem is especially devastating. ECCID should coordinate its efforts with the World Health Organization and with national structures such as the London School of Tropical Medicine in Britain and the Instituts Pasteurs d'Outre-Mer and ORSTOM in France, which have a long experience of collaborative research in developing countries.

Michel Tibayrenc

*Centre d'Etudes sur le Polymorphisme
des Microorganismes (CEPM),
UMR CNRS/ORSTOM 9926, BP 5045,
34032 Montpellier Cedex 1, France
e-mail: Michel.Tibayrenc@cepm.mpl.orstom.fr*

Straight talking

Sir—The review by the linguist David Poeppel of *The Symbolic Species: The Co-Evolution of Language and the Brain* by the neuroscientist Terrence Deacon (*Nature* 388, 734; 1997) sounds similar to a priest defending an established and powerful religion.

Where science and religion should differ is the way in which new ideas are treated.

Deacon makes a good case for a theory of evolution of language in his book, which is dismissed in Poeppel's review as merely a willingness to contribute an unimaginative opinion.

Deacon was even attacked for comparing primates with humans in his book. Unless one denies that humans evolved from primates, looking for evolutionary cues of language in non-human primates can only add to our knowledge.

After all, our language capacity does not come overnight and we are not the only species that talks and listens.

Until we fully understand the mechanisms underlying human speech and language, we should not ridicule others for suggesting alternative explanations to the long-standing problem of language evolution. The debate should not be trivialized to an argument between 'good guys' and 'bad guys'. In science, we should believe only in the truth, not in one view or the other.

**Siddhartha C. Kadia
Xiaoqin Wang**

*Laboratory of Auditory Neurophysiology,
Department of Biomedical Engineering,
Johns Hopkins University School of Medicine,
Baltimore, Maryland 21205, USA
e-mail: xwang@bme.jhu.edu*

More than just a name

Sir — You report a proposal to 'return' the name of the Royal Greenwich Observatory to the National Maritime Museum at Greenwich (*Nature* **388**, 705; 1997). We have no wish to cause offence to our parent organization, but, as with all children leaving home, there can be no going back.

The Royal Observatory moved from Greenwich in the 1950s, leaving behind a historical museum of artefacts which tell the story of UK astronomy from the foundation of the observatory by Charles II in 1675 until the middle of this century.

The Royal Greenwich Observatory at Cambridge in the 1990s is quite different. It is a modern scientific establishment with a first-class international reputation for achievements in astronomical research, telescope and instrument technology and related activities.

In our view it would be an empty gesture to detach and move the name from the observatory and return it to a museum. Furthermore, the business plan that the observatory's managers are developing to keep it going as a nonprofit organization depends crucially on the worldwide reputation that the name Royal Greenwich Observatory carries.

You suggest that the 'public

understanding' arm of the observatory might accompany the move to the museum. Our highly successful programme in this area could not be continued at a museum, and as members of the team we would not wish to try. It is successful because we not only do research ourselves but have access to experts working at the frontiers of astronomy and astronomical technology.

In his statement, the minister for science, John Battle, instructed the Particle Physics and Astronomy Research Council to "explore every avenue for keeping the institution alive". This institution is more than just a name that can be moved from one building to another.

Margaret Penston

Robin Catchpole

Royal Greenwich Observatory,

Cambridge CB3 0EZ, UK

Nonfalsifiable hypothesis

Sir — In his review of my book *Yes, We Have No Neutrons*, Walter Gratzer adopts a dismissive tone that seems to have more to do with my having chosen examples of science-gone-wrong that were not to his particular taste than with the very substantive issues involved (*Nature* **388**, 36; 1997). It is most instructive to watch how research goes off the rails when scientists

who think they have made an earth-shattering discovery forget the elements of good method. The frailties of ego undermine the discipline needed to make the very tests that will disprove the hypothesis.

In the case of the search for extraterrestrial intelligence (SETI) the hypothesis is nonfalsifiable. This does not make me a student of Karl Popper, but someone who believes that the SETI experiment is only half an experiment, at best. If aliens are sending us messages, the SETI programme is (or was) superbly equipped, we think, to pluck them from the sky. But if aliens are not sending us messages, there is no way the programme can rule out this possibility. Drake and company would merely tighten the bandwidth, add more billions of processors, search more of the sky more of the time and so on. There is no stopping rule: no test for the non-existence of such messages.

Gratzer is incorrect when he claims that receipt of such messages would prove me wrong. Although the book has a certain amount of fun with the idea of what such messages might be like, it makes no claim that they are not, in fact, being sent. It merely asserts that if no recipes for fantastic new discoveries are bathing the planet, we have no way to find out.

The review ends on a profoundly

discouraging note in which Gratzner feels compelled to quote Werfel: "For those who believe, no explanation is necessary, while for those who do not believe, no explanation is possible."

Has the general readership already hardened into two camps on each of the eight examples of "bad science" explored in the book? I think not. I hope not. Is this a general assertion about the uselessness of making any distinction between good and bad science? One would like to say "science" and "non-science" in place of these adjectives, but, as Gratzner points out, just about everything is called "science" these days.

A. K. Dewdney

Department of Computer Science,
University of Western Ontario,
London, Canada
e-mail: akd@csd.uwo.ca

Soil without life?

Sir — Over the past few weeks I have read with great interest your coverage of the discoveries of the Mars Sojourner¹. There have been many fascinating data reported about the mineral constituents of Mars rocks.

At the same time, however, references are made to martian soil without any data

to support the existence of this material. Careless use of the term soil is misleading².

For many in Earth sciences, soil imparts a notion of biological activity for which we have no evidence on Mars. In fact, most definitions of soil are closely linked with plant growth, organic matter or biological activity^{3–6}.

Although I do not subscribe to all these definitions of soil and I do not want to limit the study of pedology solely to planet Earth, I also do not envisage all martian dust as martian soil. It may well be that with continued advances in planetary studies we need to re-evaluate the definition of soil. In the meantime it may be prudent to refrain from imparting this earthly quality to all planets.

Daniel Markewitz

Woods Hole Research Center,
Woods Hole, Massachusetts 02543, USA
e-mail: dmarke@whrc.org

1. Battersby, S. *Nature* **388**, 408 & 612 (1997).
2. Johnson, D. L. *Science* **160**, 1258 (1968).
3. Dokuchaev, V. V. *Russian chernozem*. (1883). In *Selected works of V. V. Dokuchaev*, 1, 14–419 (Moscow, 1948). Israel Prog. Scient. (Transl. Jerusalem, 1967; US Department of Commerce, Springfield, VA.)
4. Byers, H. G. et al. in *Soils and Men: Yearbook of Agriculture, 1938* 979–1001 (US Department of Agriculture, Washington DC, 1938).
5. Buol, S. W., Hole F. D. & McCracken R. J. *Soil genesis and classification* 3rd ed. (Iowa State Univ. Press, Ames, IA, 1989).
6. Wilde, S. A. in *Trans. 7th Internat. Cong. Soil Sci.* **2**, 711–714 (Madison, WI, 1960).

The lost ape

Sir — Although recent correspondence has highlighted the way in which scientists are increasingly over-indulgent in their use of the concept of novelty (*Nature* **385**, 480 & **387**, 843; 1997), the practice continues.

In a recent contribution to *Nature* (**388**, 337; 1997) entitled "A new west African chimpanzee subspecies?", the authors suggest that "a previously unrecognized type of chimpanzee may be present in Nigeria and adjacent parts of Cameroon", but then go on to point out that if the subspecies is "eventually recognized, the name *vellerosus* seems to be available".

It would appear, therefore, that J. E. Gray described this potential "new" subspecies more than a century ago, a contribution that should not be overlooked simply because he lacked the technological advantages of polymerase chain reaction. Should not the recent findings be more accurately presented as a case of an old and forgotten subspecies that has been rediscovered and validated using new techniques?

Terry Harrison

Department of Anthropology,
New York University,
New York, New York 10003, USA
e-mail: harrisnt@is2.nyu.edu

IUCN's credibility critically endangered

The IUCN is the world's main authority on the conservation status of species, so it is important that its recommendations are based on sound and open science. Recent events suggest that this is not always the case.

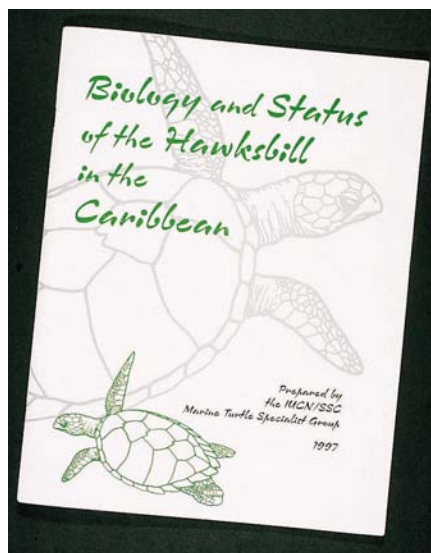
N. Mrosovsky

It is heart-warming when a politician as eminent as Norway's former prime minister Gro Harlem Brundtland writes "there is no other basis for sound political decisions than the best available scientific evidence. This is especially true in the fields of resource management and environmental protection"¹. It is thus particularly sad that the influential World Conservation Union, the International Union for the Conservation of Nature (IUCN), one of whose main aims is to provide data for scientific assessments, is not only failing to do so, but appears to be withholding information.

Although the IUCN and its main subdivision, the Species Survival Commission (SSC), do not have any legally binding authority, their opinions are considered dependable: governments, scientists, journalists and others need a quick, reliable way to find out whether particular species are in danger of extinction without having to undertake lengthy research. They often turn to the IUCN's Red Lists (formerly called the Red Data Books), which place species in various categories of risk from 'critically endangered' to 'least concern'. The IUCN also advises international conventions. At meetings of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), for instance, the IUCN produces analyses to help assess proposals affecting trade in endangered species. The IUCN's analyses are so influential because the original proposals are becoming too long and technical for national delegates to read (a recent proposal concerning sturgeon was 260 pages long), so many of them read the analyses instead.

In autumn 1996, the hawksbill sea turtle (*Eretmochelys imbricata*) was categorized in the Red Lists as critically endangered, meaning that "it is facing an extremely high risk of extinction in the wild in the immediate future". To be placed in this category, a species must meet one of various criteria, for example an 80 per cent decline in numbers over the past 10 years or three generations, whichever is the longer. To make listings objective and transparent, background information should be available. "The factors responsible for triggering the criteria, especially where inference and projection are used, should at least be logged by the evaluator, even if they cannot be included in the published lists"².

Attempts to obtain this information about the hawksbill turtle from the IUCN's



Judging a book by its cover: the controversial hawksbill document seen at the CITES meeting.

marine turtle specialist group, which made the 'critically endangered' listing recommendation, were unsuccessful; more than 9 months after the listing this information has still not been distributed.

In the meantime, however, at the tenth CITES meeting in June in Zimbabwe, Cuba proposed to 'downlist' the population of hawksbills in its waters to allow for limited trade. At this meeting, a booklet entitled *Biology and Status of the Hawksbill in the Caribbean* was used by people lobbying against the Cuban proposal. The first page of text of this booklet mentions the critically endangered listing. But nobody could have known from the booklet that the background information supporting this highly controversial status had not been made available for evaluation. In response, it was said that the booklet was not an official document, only a draft. Bound in a shiny cover with three colours (see above), it does not look like a draft. The words 'draft report' only appear inside, where they can easily be missed. The cover states 'Prepared by the IUCN/SSC Marine Turtle Specialist Group'. Many members of that group were never consulted about this document prepared in their name.

Unfortunately, this is not the only example of problems with the IUCN/SSC procedures. Some IUCN analyses, for example that concerning the Norwegian proposal for minke whales, have been criticized for containing misleading and wrong information. The analysis of the Cuban turtle proposal contained serious errors, as the authors

acknowledged in a letter to the Cuban delegation apologizing for some of these. By then, however, the damage had been done.

The IUCN makes a distinction between analyses, which do not make recommendations, and position papers, which do. In the case of the hawksbills, no position paper was ever put out. Hence the IUCN is having its cake and eating it: it distributed a document damaging to the Cuban turtle proposal, but it never came out officially against that proposal. The worst feature of the analysis was not that it contained errors, but was the secrecy surrounding some of its sources of data. Many of these are cited in the reference list as "in litt", with a name, meaning that the information is in a letter written to the IUCN by that person. The point of having reference lists is that people can look up the supporting details of statements in the text. But attempts to obtain copies of some of the letters cited in this analysis were unsuccessful. One of the letters apparently made allegations of illegal trade.

Thus the IUCN is disseminating statements derived from information that is not publicly available. What should be an analysis based on verifiable data has degenerated into assertion based on secret science. The hawksbill is listed as critically endangered, but supporting information has still not been made available. Problems with the Cuban proposal to CITES have been raised in private letters to the IUCN/SSC, and used in its analyses. No wonder some people are joking that SSC really stands for 'secret science commission'.

If it wants to regain respect, the IUCN/SSC needs to take decisive, sweeping action. Fortunately, in the case of red listings made without provision of proper documentation — and there are many examples besides sea turtles — there is a simple and scientifically appropriate solution at hand: place these species immediately in the 'data deficient' category, which can be done quickly and simply on the Internet versions of the listings. The 'data deficient' category does not imply an absence or presence of threat², but simply what it says. Such a listing might encourage those proposing a change in status to back their case by presenting data. But the data should be available at the start of the process, not added as an afterthought. □

N. Mrosovsky is in the departments of Zoology, Physiology and Psychology at the University of Toronto, Toronto M5S 3G5, Canada.

1. Brundtland, G. H. *Science* 77, 457 (1997).

2. IUCN/SSC. *IUCN Red List Categories* (IUCN, Gland, Switzerland, 1994).

Human BSE

Jeffrey Almond and John Pattison

Two sets of studies, using different approaches, provide convincing evidence that the new variant of Creutzfeldt–Jakob disease is caused by the agent that is responsible for BSE in cattle. But they cannot tell us anything about the future number of cases of this variant disease.

On 20 March 1996 the UK government announced that a distinct variant of Creutzfeldt–Jakob disease (vCJD) had occurred in ten people in the United Kingdom over the previous 14 months¹. Like CJD, the variant form results in brain damage and death; but it is pathologically and clinically distinct, not least in afflicting comparatively young people. The government also stated that the most likely cause of this apparently new disease was exposure to the agent that has caused the epidemic of bovine spongiform encephalopathy (BSE).

Since then, there have been more cases of vCJD and further evidence has accumulated that supports a link between the two diseases (Table 1). But to date there has been nothing as compelling as the results described by two papers in this issue — one, by Moira Bruce and colleagues², based in Edinburgh, appears on page 498; the other, by a group led by John Collinge (Hill *et al.*³), is on page 448. Both reinforce the conclusion that vCJD is distinct from other, sporadic forms of CJD (spCJD), and provide a convincing case that vCJD is caused by the same 'strain' of agent that has caused the BSE epidemic in Britain. In effect, vCJD is human BSE.

Both types of CJD, and other forms of transmissible spongiform encephalopathies (TSEs) such as scrapie, are thought to be caused by aberrant protein agents, which in turn cause the pathological modification of host-encoded prion proteins (PrPs). This is the 'protein only' hypothesis⁴, although it still remains possible that some other kind of foreign macromolecule is involved (see box overleaf). The two new studies^{2,3} are based on characterizing the strain of agent associated with vCJD and comparing it with strains from spCJD and from TSEs of other animals including BSE. The approaches differ slight-

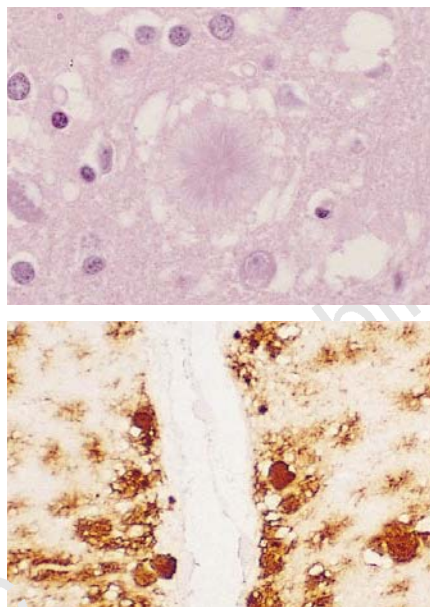


Figure 1 Tell-tale signs of vCJD. Top, cerebral cortex showing (centre) the characteristic florid plaque. Bottom, massive accumulations of PrP in the cerebellar cortex, revealed by immunocytochemistry. (Courtesy of J. W. Ironside, CJD Surveillance Unit, Edinburgh.)

ly, however: although both groups use transmission potential as a defining character, Bruce *et al.*² concentrate on incubation period and pathology, whereas Hill *et al.*³ focus on a particular feature — the so-called glycoform profiles — of the disease-specific PrP.

Evidence for the existence of multiple strains of TSE agent comes mainly from the Edinburgh group⁵ and is based on the observation that different strains will give reproducible incubation times and pathology in certain lines of inbred mice. To define strains of TSEs, they routinely use four or five mouse

lines including homozygotes for prolonged scrapie incubation periods (p7p7), for short incubation periods (s7s7) and heterozygotes (p7s7). Strains of TSEs may differ from one another both in the incubation time in a particular mouse line and in the temporal sequence in which the different lines succumb to disease. Pathological damage is measured semi-quantitatively by scoring vacuolation in nine regions of the brain and is expressed graphically as a lesion profile or 'signature'. This signature is reproducible and characteristic of a given strain⁶.

Previously, the Edinburgh group characterized eight cases of BSE from different periods in the epidemic, and from different areas: all displayed highly similar strain characteristics (their phenotype), which were moreover distinct from those observed over many years in TSEs isolated from sheep, mink and a mule deer. Importantly, strain phenotype can be maintained upon transmission between species (as assessed by re-assay in indicator mice). Thus, the BSE phenotype was maintained after experimental transmission to a pig, a goat and two sheep. Identification of the same phenotype in three cases of feline spongiform encephalopathy, an outbreak of which has occurred in the UK contemporaneously with BSE (albeit at a much lower level), and in cases in kudu and nyala in British zoos, has been taken as evidence that BSE has infected these species⁶.

Although their current research is not yet complete, Bruce and colleagues² can now also conclude the following: that the strain of agent from vCJD is: (a) the same in each of the three cases they have studied; (b) different from that seen with other forms of CJD; and (c) indistinguishable from that of BSE. The mean incubation period of vCJD in one of their inbred lines of mice (R111) is relatively short, which is characteristic of BSE whatever its source. Moreover, the signature produced in these mice is almost superimposable on that produced by BSE from cattle or from any other species accidentally or experimentally infected with BSE. The remaining lines of mice that have been inoculated with vCJD are still under observation. Experience predicts that the C57BL line of mice will be the next to develop signs of disease, and that is now happening (M. Bruce, personal communication). Given the data so far, it seems unlikely that the final results of these lengthy experiments will give a different picture — a picture which, incidentally, confirms the view that CJD occurring in farmers in recent years is in fact the sporadic form and is not related to BSE^{1,7}.

What of the work of Hill *et al.*³? Their research focuses on the fragment size and the ratio of di-, mono- and nonglycosylated forms of the disease-specific PrP protein after treatment with protease, a protein-digesting enzyme. This glycoform profile also seems largely to be maintained upon inter-species transmission, although

Table 1 Accumulated evidence of a link between BSE and vCJD

• **March 1996** Recognition of the emergence of vCJD as a predominantly UK disease (there has been one case in France), about ten years after the start of the BSE epidemic¹. Given the rarity of the appearance of new TSEs generally, emergence of a new form in the UK at this time was always unlikely to be mere coincidence.

• **June 1996** Observation¹¹ that the characteristic vCJD pathology in humans (the production of florid plaques and prominent involvement of the cerebellum; see Fig. 1) could be reproduced almost exactly in rhesus macaques by inoculation with BSE.

• **October 1996** Publication of research showing that the glycoform profile of PrP from vCJD cases was distinct from classical CJD and identical to BSE⁸.

• **October 1997** Publication of the studies by Bruce *et al.*² and Hill *et al.*³, discussed here, which reinforce the conclusion that vCJD is indeed caused by the same agent that causes BSE, and make implausible any assertion that all of these observations are coincidental.

changes may occur in certain genotypes of recipient. Last year, using this parameter, the group obtained evidence that the vCJD agent was the same as that of BSE, although its possible similarity to other animal TSEs was not ruled out⁸.

Their latest results³ include a large number of transmissions to both transgenic mice expressing the human PrP gene (*Prn-p*) and their non-transgenic counterparts, and provide stronger evidence that the agent of vCJD is distinct from those of both spCJD and the iatrogenic form of CJD (a number of cases of which were caused by treatment of patients with growth hormone derived from human pituitary glands). By most criteria, vCJD and BSE are also highly similar: their glycoform profiles are indistinguishable, in both ratios and band sizes; mice suffering from the two diseases share unusual symptoms (some of the mice walk backwards); and although details of the pathology are yet to be published, the authors refer to "striking similarities" in PrP deposition patterns. In the line of inbred mice (FVB) used for this study, there are some differences in transmission potential of BSE and vCJD. This is particularly so in the human *Prn-p* transgenic animals, to which vCJD transmits more readily than BSE. These differences are, however, probably attributable to a 'species barrier' effect for BSE but not for vCJD.

Taken together, then, the two new sets of results complement each other and give a consistent message. But can we now draw firmer conclusions about the number of cases of vCJD that will occur in the UK? Unfortunately not. To date, there have been 21 confirmed instances in the UK (each one a tragedy in its own right, and our sympathy goes out to their families). The rate of new cases is not increasing, which provides some hope that the overall number will be relatively small, but it may take several years before we can be confident that this is not a period of comparative calm before a storm. Much depends on the average incubation time of vCJD: the longer the time, the higher the final figure is likely to be⁹. At present we cannot calculate the average incubation time of BSE in humans; nor is it possible to estimate the amount of infectivity (in terms of cattle or mouse infectious doses) required to infect a human.

One observation that bears on these issues is that all of the vCJD cases examined so far are homozygous for a common amino-acid polymorphism in the human PrP protein, namely methionine at position 129. This raises the possibility that people who are homozygous for valine at this position or who are heterozygotes (about 11% and 51% of the UK population, respectively) may be relatively resistant to infection, may be subject to longer incubation times¹⁰ or may have different symptoms. The *Prn-p* transgenic mice used by Hill *et al.* have the valine 129

'Protein only' prions?

The results of Bruce *et al.*² and Hill *et al.*³ underline the need to gain a better understanding of the nature of the agents responsible for transmissible spongiform encephalopathies (TSEs). The most favoured view is that the agents are composed solely of an altered form of a host-encoded protein known as PrP and lack a foreign nucleic acid; this is the 'protein only', or prion, hypothesis⁴.

There are well-documented instances of TSE agents changing their characteristics, or phenotype, upon passage (repeated transmission through experimental animals). Indeed, Hill *et al.* report a change in fragment size upon transmission of BSE and vCJD to their transgenic mice. However, strain phenotypes can remain stable¹²; thus the Edinburgh group of

Bruce and colleagues have reported that different strains adapted to, and repeatedly passaged in, a single line of inbred mice retain their distinguishing features even though they must be composed of PrP protein with the same amino-acid sequence⁶. In addition, a single strain, as illustrated by BSE, can retain its phenotype even when passaged through different hosts. In these circumstances, then, strain phenotype is maintained even though the agent is composed of PrPs of different amino-acid sequences, raising the question of what determines strain phenotype – clearly, in these cases, it is not the primary structure of the PrP protein.

Supporters of the 'protein only' hypothesis argue that the small number of different strains that have been convincingly

demonstrated so far can be explained by different conformations of the PrP protein, perhaps in combination with modifications such as glycosylation. Such differences of course need to 'breed true' upon passage.

On the other hand, opponents of the 'protein only' hypothesis point to the potentially large number of TSE strains (the Edinburgh group claim to have evidence for around 20) and argue that an unimaginable plasticity would be required for these to be accommodated by PrP conformation alone. They suggest that a more likely explanation is that, in addition to PrP, an informational molecular component is present in the infectious agent. Whatever the nature of the agent, our understanding of TSE biology is evidently incomplete. **J.A. & J.P.**

version of the gene and it will be interesting to compare transmission to similar mice with methionine at that position.

Finally, these latest results^{3,3} also do not tell us anything more about the route by which the victims of vCJD were infected. There are various possibilities, including a common source for BSE and vCJD, and transmission from cattle to people through an intermediate species. But we still think that the most likely exposure was through eating beef products that included infected offal before it was banned from human food in late 1989. The report in the UK press of a case of vCJD in a vegetarian of 11 years' standing does not, we believe, invalidate this view; she may have inadvertently been exposed to contaminated beef products, or may have consumed infected beef before becoming a vegetarian.

A postscript. The UK government's decision, in March 1996, to point publicly to a probable link between BSE and vCJD was taken on the advice of the Spongiform Encephalopathy Advisory Committee (SEAC) — of which we were and are members, although here we are not writing in that capacity. At the time, the evidence connect-

ing the two diseases was relatively slight and SEAC's advice was rightly questioned in both the scientific and popular press. Only rarely in such circumstances can science offer definitive evidence quickly, and decisions have to depend on the weighing of uncertainties. More such judgements may yet be required concerning BSE and human disease. □

Jeffrey Almond is in the School of Animal and Microbial Sciences, University of Reading, Reading RG6 6AJ, UK. John Pattison is in the Medical School, University College London, Gower Street, London WC1E 6BT, UK.

- Will, R. G. *et al.* *Lancet* **347**, 921–925 (1996).
- Bruce, M. E. *et al.* *Nature* **389**, 498–501 (1997).
- Hill, A. F. *et al.* *Nature* **389**, 448–450 (1997).
- Prusiner, S. B. *Science* **252**, 1515–1522 (1991).
- Bruce, M. E., McConnell, I., Fraser, H. & Dickinson, A. G. *J. Gen. Virol.* **72**, 595–603 (1991).
- Bruce, M. *et al.* *Phil. Trans. R. Soc. B* **343**, 405–411 (1994).
- Hill, A. F., Will, R. G., Ironside, J. & Collinge, J. *Lancet* **350**, 188–188 (1997).
- Collinge, J., Sidle, K. C. L., Meads, J., Ironside, J. & Hill, A. F. *Nature* **383**, 685–690 (1996).
- Cousens, S. N., Vynnycky, E., Ziedler, M., Will, R. G. & Smith, P. G. *Nature* **385**, 197–198 (1997).
- Palmer, M. S., Dryden, A. J., Hughes, J. T. & Collinge, J. *Nature* **353**, 340–342 (1991).
- Lasmézas, C. I. *et al.* *Nature* **381**, 743–744 (1996).
- Bruce, M. E. *Br. Med. Bull.* **49**, 822–838 (1993).

Earth science

The puzzle of the South Pacific

Norman Sleep

Since the advent of plate-tectonic theory in the 1960s, numerous attempts have been made to find order in the complicated distribution of islands and seamounts in the South Pacific. Progress has been hindered by the remoteness of the region from oceanic institutions, and hence a lack of hard data. On page 479 of this issue¹, Marcia McNutt and her co-workers present high-resolution data from this area which, they believe, runs counter to conventional plume theory.

Most explanations of the distribution of seamounts and oceanic islands have invoked hotspot theory, where a chain of volcanoes is formed as the oceanic plate moves relative to a site of active volcanism caused by an upwelling plume from the mantle (Fig. 1a). The geometry is similar to a series of burns caused by moving your hand slowly over a candle flame, and the expected result is a regular progression in the age of seamounts and islands at increasing distance from the active hotspot.

McNutt's group did not find the expected age progression in their area, a 3° by 3° region in the Cook–Austral chain near the Macdonald seamount. Rather they find three chains of volcanic edifices of which the oldest two were previously unsuspected (see Fig. 1 on page 480). The oldest, the Ngatemato chain, is about 30 million years old, and it erupted on a seafloor that was then 10 million years old; the Taukina chain is slightly younger; and the well-known Macdonald chain has erupted in the past 5 million years.

McNutt's group was able to dredge and date only a limited number of samples. So they used an empirical correlation between gravity anomalies over seamounts and the

age of the crust at the time of seamount formation to extrapolate ages over their map area. The physics behind their method is easily understood in terms of seafloor spreading. Oceanic plates diverge at midoceanic ridge axes, and the initially hot material cools from the top down as it moves away from the ridge. The thickness of the cool, strong lid of the plate which floats on the underlying mantle increases with time.

The effect of actual loads on the plate, such as those from seamounts, can be visualized by considering a single point load. Near

the ridge where the plate is thin, the load produces a narrow region of large subsidence. Away from the ridge, a broad zone of slight subsidence is produced. The subsidence that a point load would produce is obtained mathematically from the correlation of gravity anomalies and water depth, and is expressed as an equivalent elastic plate thickness. This correlation is frozen in at the time that the seamount forms. (The situation is similar to parking a car on thin ice. A narrow downwarp forms around the car and is frozen in when the ice thickens. The downwarp does not spring back.)

In the case of the McNutt data, there are seamounts of several ages. The equivalent elastic plate obtained for the South Pacific using the actual ages of volcanism is similar to that in the other ocean basins, rather than

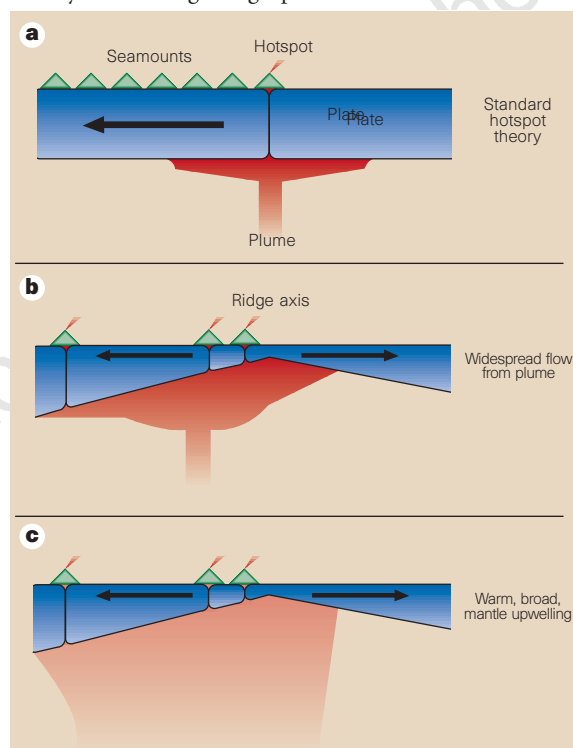
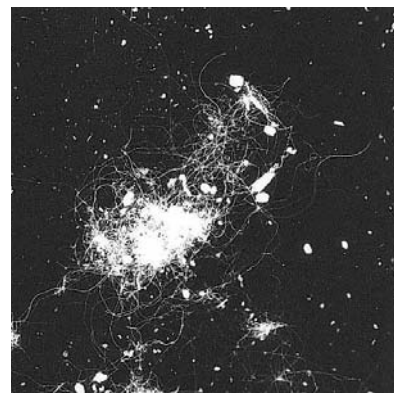
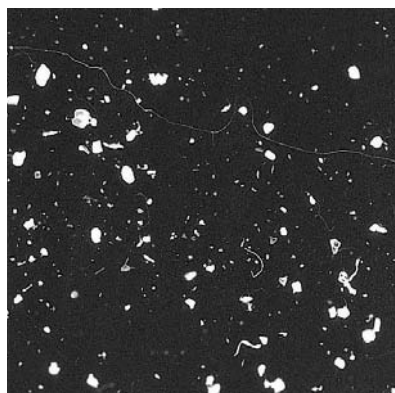
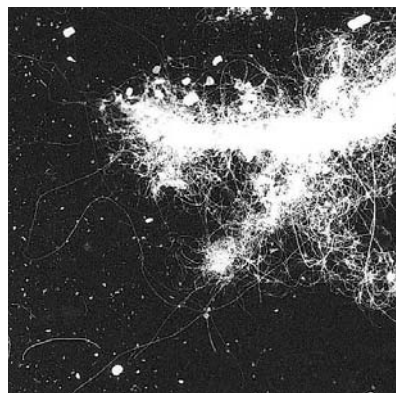


Figure 1 Standard hotspot theory (a), and variations on it (b, c), that might explain McNutt and colleagues' data¹ on the Cook–Austral chain of volcanoes in the South Pacific. c is the explanation favoured by the authors themselves. Here, a broad region of warm mantle upwelling travels beneath the lithosphere. Cracks in the oceanic plate let magma erupt to form the three chains of volcanic edifices.

Avoided object

The negatives of sound

Fluff and dust collected from the whispering-gallery, St Paul's Cathedral, London



much thinner as had previously been thought from analysis that assumed only young seamounts existed in the area.

The new information on the ages of the Ngatamoto and Taukina chains conflicts with standard hotspot theory. Nonetheless the theory provides a good explanation for the geometry of age progression of other tracks, including Hawaii, and of the volcanic activity of Yellowstone and Iceland. In the South Pacific it may be salvageable by considering the manner in which mantle plumes produce hotspots.

Mantle plumes are supposedly vertical, cylindrical conduits of hot mantle material that ascend from great depths in the Earth, probably from the core-mantle boundary. The plume is 200–300 K hotter than normal mantle, so it is buoyant relative to and less viscous than its surroundings. Lavas for hotspot volcanoes are produced by partial melting at the top of the conduit. The hot material then continues to spread laterally from the conduit along the base of the lithosphere where it may further melt causing volcanism away from the main hotspot. That is, the candle analogy should not be carried too far; plumes are sources of hot material, not point sources of heat.

The older seamount chains studied by McNutt's group were formed near the ridge axis. In this environment, buoyant plume material flows laterally along the base of the lithosphere up to the thin lithosphere at the ridge axis and is dragged away from the ridge

axis by the moving plate^{2–4}. The net effect is that a broad region of the plate is underlain by plume material and seamounts could form where melts can breach the plate (Fig. 1b). McNutt and colleagues' explanation also involves cracks in the plate that let magmas out, particularly those associated with stresses from seamount loads. However, they favour a broad region of warm mantle upwelling (Fig. 1c) rather than widespread flow from a hot plume.

How does one resolve this issue? There are questions for the dynamicist. Can broad upwelling of warm mantle and hot cylindrical plumes coexist in one convective system? If so, what does this tell us about the Earth? Can plume material from a single source explain the ill-defined tracks associated with the Austral islands? There is work, too, for the geochemical petrologist. Are the source temperatures of volcanism warm or hot and is there a pattern that helps resolve the geometry of flow? All in all, there is much room for more data. In particular, one wonders whether a detailed survey of the whole South Pacific would bring order to the complexity found near Macdonald seamount. □

Norman Sleep is in the Department of Geophysics, Stanford University, Stanford, California 94305-2215, USA.

1. McNutt, M. K., Caress, D. W., Reynolds, J., Jordahl, K. A. & Duncan, R. A. *Nature* **389**, 479–482 (1997).
2. Ribe, N. J. *Geophys. Res.* **101**, 16195–16204 (1996).
3. Sleep, N. H. J. *Geophys. Res.* **101**, 28065–28083 (1996).
4. Ito, G., Lin, J. & Gable, C. W. *J. Geophys. Res.* **102**, 15403–15417 (1997).

Evolutionary biology

Pelvic problems for mammals

Robert Presley

Living mammals can be subdivided into the monotremes (which lay eggs), the marsupials (which nurture their young in a pouch) and the placentals (in which the young stay in the uterus until a comparatively late stage of development). Traditionally, the epipubis or so-called 'mar-

supial bone' has been associated with the suckling of young in a marsupial pouch: its absence from placental mammals supposedly reflected prolonged intrauterine development and birth at a stage when the individual could survive either in a nest or by keeping pace with the mother during lactation.

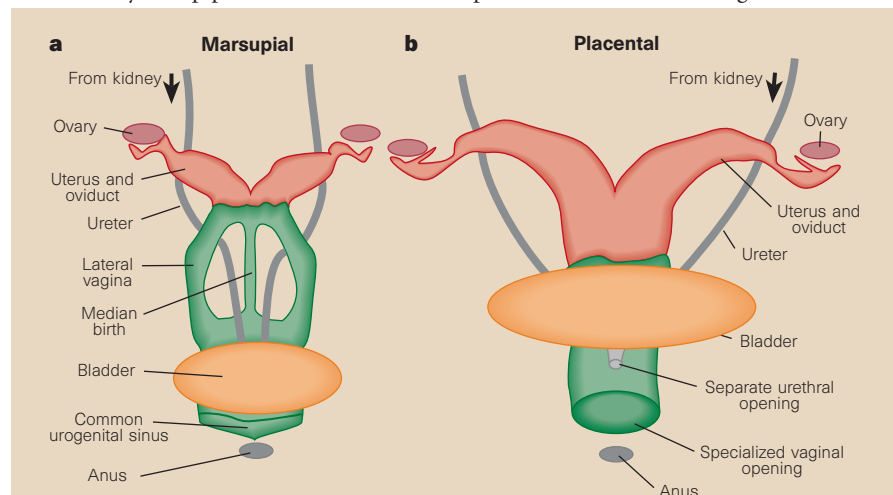


Figure 1 Ventral views of (a) marsupial and (b) placental female reproductive tracts, which develop differently in their relations at the urogenital sinus. The basic result in the marsupial is a bladder opening into a common urogenital sinus; twin vaginae lateral to the ureter (metanephric duct) as it reaches the bladder; and a narrow birth canal which is often only patent around the time of birth. The result in placentals is a separate ureteric orifice; a median patent vaginal canal; uteri opening into the vagina (the degree of fusion of the uterine canals varies considerably in placentals). These different anatomies may well have been each set up behind the cover of epipubic bones, which were once regarded as diagnostic of non-placental mammals.



Figure 2 One of the fossils described by Novacek *et al.*¹, the 80-million-year-old asioryctitherid, *Ukhaatherium nessovi*.

theory has its problems⁴: the male echidna retains its testes inside the body, and so does not have a scrotum, but it does have a pouch; in marsupials the pouch in the female probably develops from different mesenchyme than does the scrotum of males; and in those species where the female lacks a pouch the male has a scrotum.

One must also guard against climbing the *scala naturae* in considering marsupial and placental reproduction. The embryology of the developing sex ducts in relation to the ureters leads to very different anatomical relationships in marsupial and placental females (Fig. 1). The marsupial has bilateral vaginæ and a narrow, soft-tissue birth canal. It is difficult to visualize how, early in the evolution of placental mammals, diverse groups could have started from the marsupial pattern and arrived independently at the quite different plan found throughout present placentals.

Surprisingly little is known about epipubic bones in living mammals. Notwithstanding major differences of locomotor habit in marsupials, these bones are present in species belonging to all marsupial lineages. They are paired, with a plane-faced synovial joint connecting each to the cranial part of the body of its pubis flanking the symphysis. Each bone points cranially as a rod or narrow blade with its long axis lying in the muscle of the abdominal wall. In a platypus specimen of mine, the adductor longus muscle of the leg extends onto this part of the abdomen from the adjacent pubic tubercle. So there must be interaction between trunk and leg muscles through the epipubis.

No detailed anatomy of the connections and function of epipubic bones exists for most marsupial species: an extensive survey⁵ both of specimens and of the available literature concluded that, while the jury awaits much more evidence, the balance so far is that they have musculo-skeletal functions—probably, they serve to stiffen the median part of the ventral abdominal wall. As is painfully known to hernia sufferers, in mammals the abdominal wall is weakened by the inguinal canal which allows the seasonal or

permanent dropping of the testes, and by associated ducts and cooling rete of blood vessels. There is an equivalent, though lesser, weakening caused by the descent tract in the female, even though the ovary does not move down it. So we must consider more than mammary tissue and pouches when these bones occur.

The challenge highlighted by these new fossil findings¹ is that we do not know what any possessor of the epipubis does with it, let alone what all possessors do with it. It may be an element in the evolutionary history of mammalian reproductive habits, or it may

have something to do with the mechanics of the lower trunk and abdomen. It is part of our heritage that cannot simply be dismissed as a bit of bone which proves that something is a marsupial. □

Robert Presley is in the School of Molecular and Medical Biosciences, Cardiff University of Wales, PO Box 911, Cardiff CF1 3US, UK.

1. Novacek, M. J. *et al.* *Nature* **389**, 483–486 (1997).
2. McCrady, E. *Am. Anat. Mem.* **16**, 1–233 (1938).
3. Bolliger, A. *Proc. R. Soc. New South Wales* **76**, 283–293 (1943).
4. Renfree, M. B. in *Mammal Phylogeny* Vol. 1 (eds Szalay, F. S., Novacek, M. J. & McKenna, M. C.) 4–20 (Springer, Berlin, 1993).
5. White, T. D. *J. Theor. Biol.* **139**, 342–357 (1989).

Kinetic theory

Optimal paths and irreversibility

Mark Millonas

What do mutations in a genetic sequence, conformational changes in proteins, an electron escaping out of a deep energy well, and the nucleation of galaxies in the early Universe have in common? They are all events that involve large, rare fluctuations. These 'stellar' events are a link between the violent fluctuations of the microscopic world of particles and fields and the ordered world of cells and computer chips. Luchinsky and McClintock (page 463 of this issue¹) have directly studied these large, rare events in the laboratory, and provide an experimental illustration of the asymmetry between the path away from an average and the path back (Fig. 1).

The task of producing a theory for understanding fluctuating nonequilibrium systems, one comparable in scope to that developed over the past century or so for equilibrium systems, is beset by a number of funda-

mental difficulties. An equilibrium system such as a chemical reaction coupled to an heat bath is constrained by the symmetry of microscopic reversibility. The strangeness of a nonequilibrium system such as a chemical reaction driven by an external high-grade energy source arises because it is not constrained by this symmetry. As a consequence, every time a fluctuation occurs the system might pick up or lose energy in a complicated way, leading to many types of unusual behaviour.

The theory of large fluctuations provides a possible way out of some of these difficulties because it is a way of treating just the very large fluctuations. Although uncommon, such fluctuations often control important types of macroscopic behaviour. For example, the conformational change of an enzyme is a relatively rare event that can lead to a complex cascade of biochemical reactions in a cell. Similarly, the rare escape of electrons from an energy well can determine most of the electrical properties of a semiconductor.

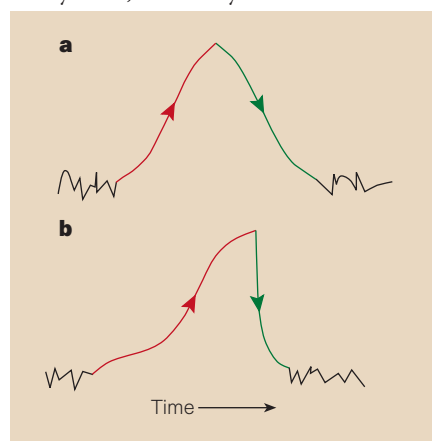


Figure 1 Asymmetry of large fluctuations in irreversible systems. a, In equilibrium systems, each large fluctuation (red) away from an average must be matched by a mirror-image return to the average (green). b, In non-equilibrium systems, the fluctuations are instead irreversible. Luchinsky and McClintock¹ have observed this theoretically predicted asymmetry in simulations using electrical circuits.

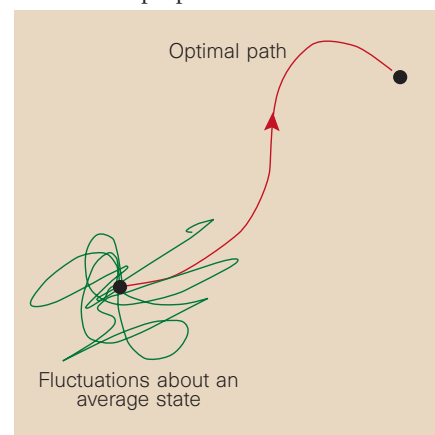


Figure 2 The optimal path. There are many paths of roughly the same probability that reach points near the average (green), but only one dominant path that reaches each point far from the average (red). These paths often determine important types of nonequilibrium behaviour.

To understand the way the theory of large fluctuations works, one must picture all the possibilities of a system's motion. These can be represented as paths in a mathematical phase space (Fig. 2). Each path the system might take can be assigned a particular probability density, related to a mathematical quantity called the 'action'. The probability decreases very sharply as the action increases, and the rate of occurrence of very rare events is determined not by all possible paths, but with overwhelming weight by the most probable, or optimal path, and a narrow 'tube' of trajectories about it. Thus many important problems can be reduced to ones involving an analysis of just one dominant path. Luchinsky and McClintock's study¹ is an important step in assessing the reliability of these theoretical ideas to real-world situations.

The authors were able experimentally to study models originally suggested by theorists to illustrate several types of nonequilibrium behaviour. What made this possible was the clever trick of doing electrical analogue simulations of the systems. In this way the relationships between the theoretical objects and experiment can be demonstrated very cleanly. However, data from 'real' experiments may still be a somewhat remote prospect. Although the optimal paths are

observable, Luchinsky and McClintock's simulations sometimes required data to be collected for as long as a week.

Whereas the behaviour of equilibrium systems is 'smooth', large-fluctuation treatments of nonequilibrium systems^{2,3} give rise to singularities — places where behaviour changes suddenly. Opinions varied as to what these singularities signified, and some researchers asserted that the method of large fluctuations simply was not valid in such cases. But in recent years a topological analysis of optimal paths (Fig. 3) has proved that the physically important singularities can be understood within the framework of the theory of large fluctuations⁴.

What excites researchers is the wealth of new and unexpected phenomena related to these singularities. An example is the way the optimal path in a system such as a nonequilibrium chemical reaction can suddenly change shape⁴⁻⁶, leading to a 'kink' in the plot of reaction rate versus some parameter^{5,6}. This is a result that is sharply at odds with the 'escape rates' of a system in equilibrium⁷, the basis for understanding chemical reactions and similar phenomena for the past half century.

This emerging topological viewpoint is analogous to the viewpoint of chaos in nonlinear dynamical systems. As Poincaré

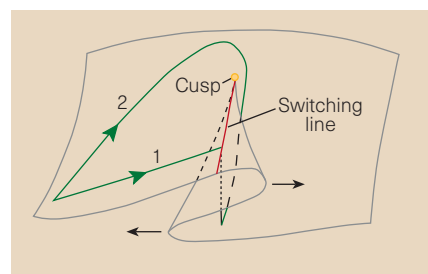


Figure 3 Switching in nonequilibrium systems. As some physical parameter is varied, the sheet on which the optimal path lies (in an abstract mathematical space) can fold over. The further from equilibrium the system gets, the greater the chance that such folds will develop. The optimal path to go from one side of the switching line is direct (path 1), whereas the optimal path to go from the other side of the switching line must detour all the way around the cusp point of the fold (path 2). The switching of optimal paths that arises with the appearance of the fold can lead to a sudden change in the behaviour of the system.

realized near the turn of the century, the dynamics of nonlinear systems are often far too complicated to be comprehensible by ordinary analysis. He suggested that the way to understand such systems was through a study of the geometric patterns of their

Avian sex-ratios

Pretty Polly, Polly, Polly, Polly, Polly ...

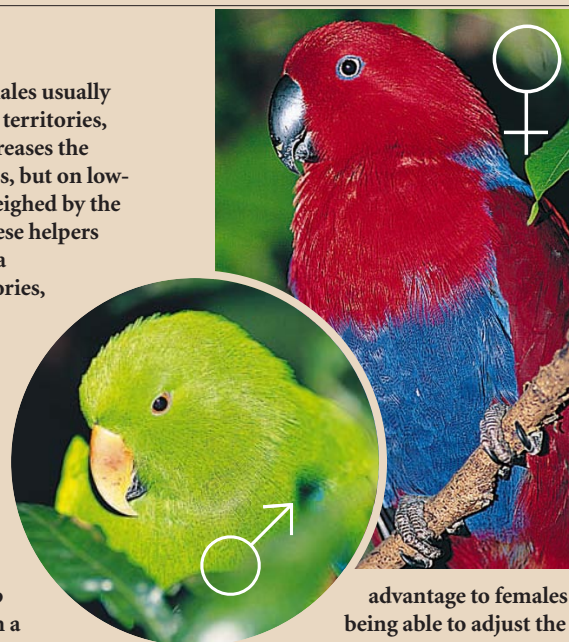
Most organisms produce roughly equal numbers of male and female offspring and, many years ago, R. A. Fisher proposed the explanation for this — in a population in which one sex predominates, the members of that sex will produce fewer offspring (on average) than will those of the minority sex. So, genotypes that produce more of the minority sex than the population average will be favoured.

There is increasing evidence that some female birds control the ratio of sons to daughters in their families^{1,2}. The latest example is reported by Heinsen *et al.*³ in this month's *Proceedings of the Royal Society*. The authors found that most eclectus parrots (*Eclectus roratus*) produce long runs of offspring of one sex — one female produced 20 sons in succession, followed by a run of 13 daughters.

Avian sex-ratios may depend on the position of an egg in the laying sequence, the date of breeding, maternal age, and the sexual attractiveness and condition of the father². Before the new study, the most extreme variation in sex allocation in birds was shown by the Seychelles warbler, *Acrocephalus sechellensis*⁴. In this species, young females often remain on their parents' territories and help to rear subsequent offspring (cooperative

breeding), whereas young males usually move away. On high-quality territories, help from the daughters increases the parents' reproductive success, but on low-quality territories it is outweighed by the amount of resources that these helpers consume. So, daughters are a disadvantage on poor territories, and only 13 per cent of offspring produced on such territories are female, contrasting with 77 per cent on good territories.

The Seychelles' warbler story is relevant to the new study because — uniquely among parrots — *E. roratus* is also a cooperative breeder, with up to ten males associating with a single breeding female (or, sometimes, with one or two other females). Unfortunately, the natural history of the species is poorly known. Moreover, the sex-ratio results were based on the records of aviculturalists, who keep the birds in traditional pairs. These are so different from the natural breeding system that they do not provide clues as to what determines sex-ratio adjustment. One can only speculate that, in nature, there is an



advantage to females in being able to adjust the sex of their offspring in relation to the relative numbers of males and females in the breeding group.

Jeremy J. D. Greenwood

Jeremy J. D. Greenwood is at the British Trust for Ornithology, Thetford, Norfolk IP24 2PU, UK.

1. Frank, S. A. *Annu. Rev. Ecol. Syst.* **21**, 13–55 (1990).
2. Ellegren, H. & Sheldon, B. C. *TREE* **12**, 255–259 (1997).
3. Heinsen, R., Legge, S. & Barry, S. *Proc. R. Soc. Lond. B* **264**, 1325–1329 (1997).
4. Komdeur, J., Daan, S., Tinbergen, J. & Mateman, C. *Nature* **385**, 522–524 (1997).

dynamics. Just as this geometric pattern helps us think about the erratic, unpredictable dynamics of deterministic non-linear systems even without being able to perfectly predict their behaviour, the geometrical approach to large fluctuations allows us to anticipate the broad characteristics of nonequilibrium systems.

One of the conundrums that emerges on the mysterious threshold between the microscopic and macroscopic worlds is the arrow of time. Irreversibility arises when we make the jump from the reversible, microscopic laws of classical and quantum mechanics to descriptions where the collective action of a complex external world is modelled as a source of fluctuations and dissipation. The theory of large fluctuations will not tell us anything new about the fundamental mean-

ing of irreversibility, because its axioms start on an irreversible level; but it can answer many important questions about how non-equilibrium fluctuations lead to complex macroscopic behaviour. □

Mark Millonas is in the Department of Pharmacological and Physiological Sciences, University of Chicago (MC6094), 5841 South Maryland Avenue, Chicago, Illinois 60637, USA.

1. Luchinsky, D. G. & McClintock, P. V. E. *Nature* **389**, 463–466 (1997).
2. Freidlin, M. I. & Ventzell, A. D. *Random Perturbations of Dynamical Systems* (Springer, New York, 1984).
3. Graham, R. in *Noise in Nonlinear Dynamical Systems* Vol. 1 (eds Moss, F. & McClintock, P. V. E.) 225–278 (Cambridge Univ. Press, 1989).
4. Dykman, M. I., Millonas, M. M. & Smelyanskiy, V. N. *Phys. Lett. A* **193**, 53–58 (1994).
5. Dykman, M. I. *Phys. Rev. A* **42**, 2020–2029 (1990).
6. Maier, R. S. & Stein, D. S. *Phys. Rev. E* **48**, 931–938 (1993).
7. Kramers, H. A. *Physica* **7**, 284–304 (1940).

Molecular clocks

PERpetuating the PAST

Paolo Sassone-Corsi

The rhythmicity of biological processes is governed by intracellular mechanisms — a molecular time-keeping organizer seems to regulate the balance of many periodic physiological events, ranging from plant photosynthesis to hormonal oscillations and annual breeding cycles in mammals. A major property of this endogenous clock, or pacemaker, is its self-sustained rhythmic function. Pacemakers are thought to be made up of a network of transcriptional regulators, which mediate physiological events through the cyclic control of other genes downstream.

Studies in various organisms indicated that the regulatory mechanisms that are responsible for clock function seem to share some limited similarities. So, considering that the day–night rhythm has remained unchanged during the past hundreds of million of years, one question arises: how conserved are the molecular features of the clock during evolution? The cloning of the mammalian orthologue of the *Drosophila period* (*per*) gene by Tei and colleagues¹, reported on page 512 of this issue, represents another step towards answering this question.

The fruitfly *per* gene is essential for rhythmic adult eclosion (emergence from a pupa-case) and locomotor activity. It has been considered as a model for clock genes^{2,3}, and many laboratories have tried to identify *per* homologues in mammals. But, until now, *per* genes had been cloned only in dipterans, with the exception of the giant silkworm *Antheraea pernyi*, which diverges from *Drosophila* by 240 million years⁴. Steven Reppert and colleagues have shown that although there is considerable conservation between the structural features of the silkworm and fruitfly *per* genes, they seem to

elicit the clock function by completely distinct mechanisms^{4,5}.

Now the comparison can be extended to mammals. Tei and colleagues have successfully applied a sophisticated polymerase-chain-reaction cloning approach to fish out the mouse and human *per* genes. They have based their quest on the idea that a region of the dPer protein — the so-called PAS domain — would be the most likely structural feature to be conserved during evolution. PAS domains were named after the three proteins in which this structural motif was first identified: the *Drosophila* Per, the mammalian Arnt (a dimerization partner of the dioxin receptor) and Sim, which is the product of the fruitfly *single-minded* gene.

PAS domains are around 260 amino acids in length, and they contain two direct repeats

of about 60 amino acids⁶. They are now known to be common to several transcription factors, and they are often coupled within the protein to a basic helix–loop–helix (bHLH) domain. Importantly, both fly and mammalian Per proteins lack a well-defined bHLH domain, suggesting that they may act as transcriptional regulators but not necessarily as DNA-binding proteins. Indeed, PAS has been shown to act as a dimerization interface in several cases and to determine, depending on the association partner, either positive or negative transcriptional regulation^{7,8}. The PAS domains of various bHLH/PAS *Drosophila* proteins also confer target-gene specificity⁹.

The PAS domain is generally represented as PAS-A and PAS-B subdomains, because the region of homology is divided by a stretch of non-conserved amino acids. The identity between the fruitfly and mouse Per within the PAS domains is strikingly high (Fig. 1), suggesting that the two proteins could share structural and functional properties. One essential observation made by Tei and colleagues¹ is that *per* expression oscillates in the hypothalamic suprachiasmatic nucleus — the location of the endogenous clock. So, as with the fly gene, *per* in mammals could work directly as a molecular element of the clock.

The conservation of the PAS domains begs the question: what could be the partners for Per in the regulation of clock functions? In the fruitfly, Per dimerizes with Tim (the product of the *Drosophila* gene *timeless*, *tim*), the levels of which are also regulated by a feedback loop. The Per and Tim loops are interdependent, and they control the nuclear translocation of the Per–Tim heterodimer^{10–12}. The timing of nuclear entry of the Per–Tim complex probably directly modulates clock function in the fruitfly. The obvious question of whether Per dimerizes with Tim in mammalian cells cannot be

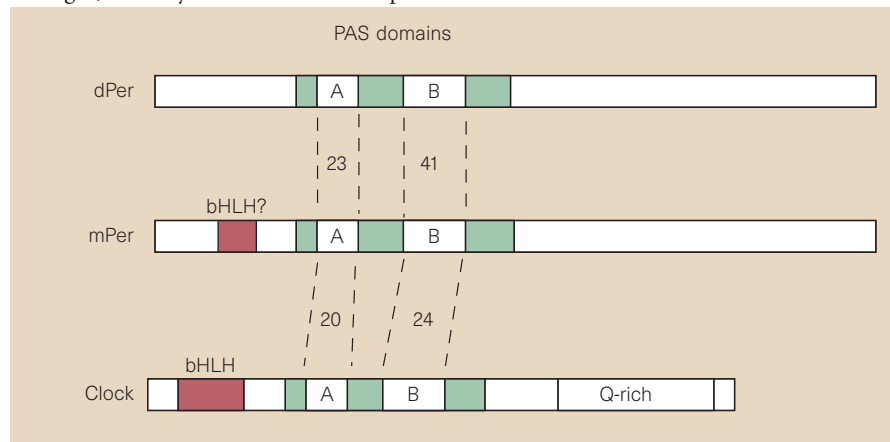


Figure 1 Comparison of the fruitfly and mouse Per proteins, along with the mouse Clock protein. Tei *et al.*¹ have cloned the mouse and human *per* genes, and they show that the product of the mouse gene oscillates in a circadian manner, in the same way as *Drosophila* Per. Moreover, there is significant sequence identity between the PAS domains (PAS-A and PAS-B) of these proteins and the mouse *clock* gene product. PAS domains are involved in protein–protein interactions, so distinct clock proteins may cooperate to determine molecular rhythmicity.



100 YEARS AGO

It is much to be regretted that the splendid male Giraffe, presented to the Queen by the Chief Bethoen, of Bechuanaland, died so soon after reaching this country. It was placed on board the s.s. *Roslin Castle* ... and left Cape Town on September 1. The passage was a stormy one, and after the first week the Giraffe declined to eat anything but bread. It was therefore, on reaching London, nearly dead from exhaustion, and only lived for half an hour after its arrival at Regent's Park. This event is the more to be lamented, as the fine young female already in the Society's Gardens ... thus remains without a mate, and there is at present little prospect of obtaining one.

Prof R. C. Carpenter, professor of engineering science in Cornell University has (says *Engineering*) been conducting an elaborate set of experiments on bicycle friction which have led him to the conclusion that no form of gearing can possibly equal the best chain for efficiency and durability. With such the frictional loss has been found to be between $\frac{1}{2}$ and $\frac{3}{4}$ per cent. of the total power transmitted, this result being obtained with a chain which had previously been ridden more than 2000 miles with a rider weighing about 14 stone. With some other chains less well constructed, a greater loss has been found, the friction lying generally between 2 and 5 per cent.; the maximum shown, even by an old chain which did not fit its sprocket properly, was under 10 per cent. No bevel gears as yet constructed give as good results as these. From *Nature* 30 September 1897.

50 YEARS AGO

A U.S. Air Force "Skymaster" transport aircraft flew from Stephenville, Newfoundland, to Brize Norton R.A.F. Station in Oxfordshire, on September 22, a distance of 2,400 miles across the Atlantic, including taking off and landing, without being controlled in any way by the crew on board. ... The apparatus automatically controls the take-off and climb to an arranged height. It then homes on a beam sent out by a radio beacon. ... Finally, it sets the machine into the required glide, lands and brakes. In the United States the system is understandably described as 'push button flying'. From *Nature* 4 October 1947.

answered just yet, because the mammalian *tim* homologue has still not been identified. But the search for Per partners in mouse or human cells will now begin, most likely by using the PAS domain as a bait.

A molecule that could potentially interact with Per in the mouse is the product of the *clock* gene¹³, which has been identified by positional cloning in the laboratory of Joseph Takahashi. Mice carrying a single mutation in the Clock protein have a dramatically altered circadian rhythmicity. Strikingly, Clock also contains a PAS domain, which shows significant homology to the fruitfly and mouse Per proteins (Fig. 1). But the Clock protein also contains a bHLH domain, so, in a hypothetical model of Per–Clock interaction, Per could modulate the activity of Clock. Other possibilities may include the association with additional cofactors, or many combinatorial interactions with other, as-yet unidentified, Per- and Clock-like molecules. Importantly, a mouse gene (*RIGUI*) with homology to dPer has just been cloned¹⁴. Although it is likely that mPer and RIGUI are products of the same gene, RIGUI is thought to contain a putative bHLH domain (Fig. 1) and to oscillate in the retina.

Circadian clocks are common to most living organisms, and they are the best example of physiological adaptation. So this is a very exciting time for those working in the fields of biological rhythms and gene expression — the collection of tools that are necessary for an understanding of the molecular clock is now, dramatically, getting richer. □

Paolo Sassone-Corsi is at the Institut de Génétique et de Biologie Moléculaire et Cellulaire, BP163, CU de Strasbourg, 67404 Illkirch, France. e-mail: paolosc@igbmc.u-strasbg.fr

1. Tei, H. *et al. Nature* **389**, 512–516 (1997).
2. Konopka, R. J. & Benzer, S. *Proc. Natl Acad. Sci. USA* **68**, 2112–2116 (1971).
3. Hall, J. C. *Trends Neurosci.* **18**, 230–240 (1995).
4. Sauman, I. & Reppert, S. M. *Neuron* **17**, 889–900 (1996).
5. Sassone-Corsi, P. *Nature* **384**, 613–614 (1996).
6. Nambu, J. R., Lewis, J. O., Wharton, K. A. & Crews, S. T. *Cell* **67**, 1157–1167 (1991).
7. Huang, Z. J., Edery, I. & Rosbash, M. *Nature* **364**, 259–262 (1993).
8. Lindeboom, M. C., Poellinger, L. & Whitelaw, M. L. *EMBO J.* **14**, 3528–3539 (1995).
9. Zelzer, E., Wapner, P. & Shilo, B.-Z. *Genes Dev.* **11**, 2079–2089 (1997).
10. Seghal, A. *et al. Science* **207**, 808–810 (1995).
11. Zeng, H., Qian, Z., Myers, M. & Rosbash, M. *Nature* **380**, 129–135 (1996).
12. Vosshall, L. B., Price, J. L., Seghal, A., Saez, L. & Young, M. *Science* **263**, 1606–1609 (1996).
13. King, D. P. *et al. Cell* **89**, 641–653 (1997).
14. Sun, Z. S. *et al. Cell* **90**, 1003–1011 (1997).

Oceanography

Shades of the sea

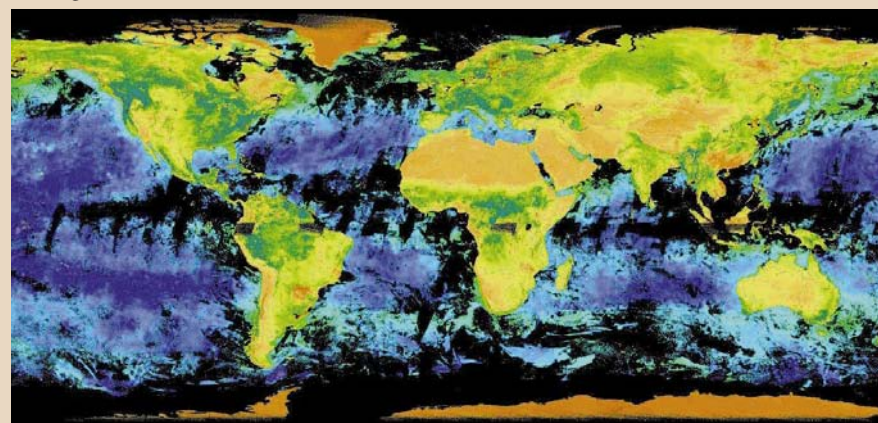
Last month the first results emerged from the long-awaited remote Earth-sensing instrument, SeaWiFS (the Sea-viewing Wide Field-of-view Sensor), which is designed to monitor the colour of the oceans. Ocean colour data provide a means to investigate the abundance of phytoplankton, suspended sediments and organic materials in the surface ocean.

The viability of using satellite-based instruments to measure ocean colour was established over a decade ago. But although the Coastal Zone Color Scanner (a forerunner to SeaWiFS, which operated between 1978 and 1986) provided some information on biological productivity, the data lacked crucial spatial and temporal coverage.

SeaWiFS boasts global coverage every two days, higher sensitivity to visible light, and an improved capability to 'see' through the atmosphere. Moreover, some of the SeaWiFS images (as shown here) convey land-vegetation information as well as data on oceanic chlorophyll concentrations (which are lowest in the purple regions, and higher in the green sea areas).

Studies of ocean colour also have potentially wider social and economic implications. The ability to find pockets of high phytoplankton density, for example, could help in the management of fish stocks. Moreover, SeaWiFS can keep a watchful eye out for harmful blooms of coastal algae, which have been linked to outbreaks of human disease.

Karen Southwell



NASA/GSFC

Palaeoanthropology

One skull does not a species make

Eric Delson

The palaeoanthropology community must look quite Pavlovian to outsiders — we all drool predictably every time a new fossil is discovered. In fact, such new specimens may be important not only in their own right, but in terms of wider implications for the field. Such a fossil from Konso in Ethiopia is described on page 489 of this issue by Suwa and colleagues¹, almost 40 years after Mary Leakey found, and Louis Leakey described², the first known member of its species.

Australopithecus (or *Paranthropus*) *boisei* is the morphologically most specialized member of an extinct early human clade (branch), which separated from the lineage that eventually led to living people some 3–2.5 million years ago (Myr). These so-called robust australopithecids were characterized by extremely large and thickly enamelled molar teeth and jaws, which were moved by massive chewing muscles attached

to a sagittal crest (fore-and-aft ridge of bone) on top of the relatively small-brained cranium. In the two later species of this group — East African *P. boisei* and South African *P. robustus* — the incisors and canines were strongly reduced, and were even smaller than our own.

As with many early human species, *P. boisei* occurs at only a few sites: Olduvai Gorge and Peninj (Tanzania), Chesowanja (Kenya) and, especially, in the Lake Turkana Basin (in the Shungura, Koobi Fora and Nachukui/West Turkana zones on the Kenya–Ethiopia border), mainly between 2.2–1.6 Myr^{3,4}. A partial upper jaw has also been recovered from the Chiwondo Formation in Malawi, by Tim Bromage, Friedemann Schrenck and colleagues.

The discovery of a *P. boisei* skull by Suwa *et al.*¹ now extends the geographical range to the northeast (Fig. 1). Temporally, with a firm date of 1.4 Myr, it is near the previously

extrapolated latest occurrence of the species. Most of the other finds are associated with relatively moist and wooded habitats, but the Konso deposits yielding the new specimens indicate a more open, drier palaeoenvironment (although nearby subsites were also moist and wooded). In the same bed at Konso⁵ are a mandible of *Homo erectus* and an assemblage of artefacts referred to the earliest Acheulean (hand-axe) industry. In fact, revised calibration of the Olduvai Gorge sequence^{4,6} indicates that the oldest Acheulean sites there (EF-HR and MNK; see ref. 7) probably date to 1.6–1.5 Myr.

The new specimen KGA10-525 is one of the most complete skulls of *P. boisei* yet known, comprising much of the cranium and part of the lower jaw — the first association of these two elements in a single, well-preserved individual of this species. Future studies may provide new insight into the biomechanics of mastication in *P. boisei*, given the potential to look at the opposing upper and lower jaws. But perhaps the greatest importance of this specimen is its implication for the study of variation in this, and other, human populations and species.

During the 1960s there was much discussion as to whether *P. robustus* and *P. boisei* were distinct species or merely geographical variants of a single form. After Tobias' monographic treatment of the original Olduvai material⁸ and the recovery of many specimens from the Turkana Basin, recognition of two species was common. Rak's delineation⁹ of facial differences between the known samples supported this view, and increasing use of the generic name *Paranthropus* reflects a growing acceptance of Robinson's arguments¹⁰ that these two species (and their conservative relative *P. aethiopicus*) formed a distinct clade in human evolution. Nonetheless, it has been clear that the Olduvai type specimen¹ is the most 'extreme' example of *P. boisei*, whereas some of the crania from Turkana are — in certain features — morphologically intermediate between this specimen and the generally smaller and less robust South African population.

The cranium discovered by Suwa *et al.*¹ is large, even for *P. boisei*, and it has many features that place it within that species. But it also lacks some of the diagnostic features, such as the 'visor-like' cheekbones, the forward surface of which projects obliquely out and down — like a knight's faceplate — perhaps to take up chewing strain. Moreover, the sagittal crest is emphasized posteriorly, almost as in *P. aethiopicus*, and the palate is broad and short, similar to the shape seen in *Homo* (see Table 2 of the paper, page 491). Finding these atypical character states in *P. boisei* from a new population complements the



Figure 1 Map showing some of the East African localities that have yielded fossils of *Paranthropus*. All have *P. boisei*; *P. aethiopicus* also occurs at Nachukui, Omo and Koobi Fora. The skull described by Suwa *et al.*¹ comes from Konso — the most northeasterly and probably the youngest of these occurrences.

observations of Brown *et al.*¹¹ on variability in crania of this species within the Turkana Basin.

Some researchers might 'solve' the inconsistency by naming a new species for the Konso population, but both the authors and I would caution against such an oversimplified approach. Neighbouring species of living animals are generally distinguished by a lack of gene exchange (interbreeding and mate-recognition). But this biological species concept (BSC) cannot be directly applied to modern animals that are separated in space or to fossils. In both of these cases, some specialists try to draw analogies with the BSC by studying the degree of variation and overlap among well-studied neighbour species. Others argue that any discernible difference in morphology implies that interbreeding would have been unlikely, thus according full species rank more readily (the phylogenetic species concept; PSC).

The sister taxa *P. boisei* and *P. robustus* are probably the most similar species pair within the Hominini (post-ape human lineage). The new sample from Konso opens, once again, the question of their taxonomic distinction — if the diagnostic features of both species co-occur in a single population sample, are they really different species? Only by re-evaluating character-state distributions across the robust australopithecines can we hope to answer this question.

Looking at an even broader scale, the number of widely accepted species in our genus, *Homo*, has increased steadily, stemming in part from the PSC philosophy presented in a seminal essay by Tattersall¹². Populations spanning the last 1 Myr were, for many decades, considered varieties of our own species *H. sapiens*. But the Neanderthals of western Eurasia (dating roughly from 200,000–30,000 yr) are now thought by many to represent a distinct species *H. neanderthalensis*, which had its roots in earlier European populations genetically isolated from African contemporaries.

The PSC advocates have recognized as many as three additional species in the interval between 1 Myr and 200,000 yr: *H. rhodesiensis* for the African populations, which are younger and more derived than *H. erectus*; *H. heidelbergensis* for European groups in the 500,000–200,000-yr range; and, most recently, *H. antecessor* for a European sample dating around 800,000 yr. *Homo antecessor* is said to show facial features that are reminiscent of *H. erectus* and some later populations, leading to the suggestion¹³ that it is the common ancestor of all of the younger species. But that interpretation is based mainly on the face of a single subadult individual, with no idea of the variation (adult or juvenile) in that sample or in juveniles of either *H. heidelbergensis* or *H. rhodesiensis*.

As Suwa *et al.*¹ note, the Konso specimens underline the importance of understanding

intraspecific variation before erecting new species based on single specimens or populations. Alternative taxonomies must be carefully considered against comparisons of modern and fossil variation. For example, it is still reasonable to include all Middle and Late Pleistocene *Homo* in *H. sapiens*, perhaps differentiated as spatio-temporal subspecies. Alternatively, the Neanderthal lineage in Europe and southwest Asia might be classed as a single species *H. neanderthalensis* (including *H. heidelbergensis* and, perhaps, *H. antecessor* as temporal stages). In such a model, the role of '*H. rhodesiensis*' is unclear, in part depending on whether it includes populations that were, ultimately, ancestral to modern *H. sapiens*, as the 'Out of Africa' hypothesis implies.

Even with the discovery of Neanderthal genetic material¹⁴, we still cannot decide whether the Neanderthals were one of several related species in an extinct radiation, a single species close to our own, or a 'race' of *H. sapiens* (with that species redefined to include 2-Myr-old *H. erectus*). How are we to choose among these or other alternatives? One approach might be to compare variation within and among the several named samples to that found in modern humans (either quantitatively or in terms of character distribution), under the hypothesis that geographical variation today is broadly comparable to past spatio-temporal variation¹⁵.

The Konso sequence has been carefully prospected for the past six years by a truly international team, with primary researchers from Japan, Ethiopia and the United States. It is perhaps fitting that a group which is so representative of modern human diversity should discover a fossil that helps us to understand the origins of that diversity. □

Eric Delson is at the American Museum of Natural History, New York, New York 10024, and in the Department of Anthropology, Lehman College, Bronx, New York 10468, USA.

1. Suwa, G. *et al.* *Nature* **389**, 489–492 (1997).
2. Leakey, L. S. B. *Nature* **184**, 491–493 (1959).
3. Suwa, G., White, T. D. & Howell, F. C. *Am. J. Phys. Anthropol.* **101**, 247–282 (1996).
4. Kimbel, W. H. in *Paleoclimate and Evolution, with Emphasis on Human Origins* (eds Vrba, E. S., Denton, G. H., Partridge, T. C. & Burckle, L. H.) 425–437 (Yale Univ. Press, New Haven, 1995).
5. Asfaw, B. *et al.* *Nature* **360**, 732–735 (1992).
6. Tamrat, E., Thouveny, N., Taieb, M. & Opdyke, N. D. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **114**, 273–283 (1995).
7. Leakey, M. D. *Olduvai Gorge Vol. 3: Excavations in Beds I and II 1960–1963* (Cambridge Univ. Press, 1971).
8. Tobias, P. V. *Olduvai Gorge Vol. 2: The Cranium and Maxillary Dentition of Australopithecus (Zinjanthropus) boisei* (Cambridge Univ. Press, 1967).
9. Rak, Y. *The Australopithecine Face* (Academic, New York, 1983).
10. Robinson, J. T. in *Evolutionary Biology Vol. 1* (eds Dobzhansky, T., Hecht, M. K. & Steere, W. C.) 69–100 (Appleton-Century-Crofts, New York, 1967).
11. Brown, B., Walker, A., Ward, C. V. & Leakey, R. E. *Am. J. Phys. Anthropol.* **91**, 137–159 (1993).
12. Tattersall, I. *J. Hum. Evol.* **15**, 165–175 (1986).
13. Bermúdez de Castro, J. M. *et al.* *Science* **276**, 1392–1395 (1997).
14. Krings, M. *et al.* *Cell* **90**, 19–30 (1997).
15. Delson, E. in *Evolutionary Biology at the Crossroads* (ed. Hecht, M. K.) 141–145 (Queens College Press, New York, 1990).

Daedalus

The ears have it

What makes an animal an appealing pet? One requirement is a face that seems 'human'. Cats, dogs and other popular pets have faces reminiscent of human infants, and trigger protective emotions in us. They also convey emotions by signals which we ourselves recognize. Daedalus now points out that the converse is also true. Good pets must be able to react to us, and to interpret our expressions and behaviour correctly. This is a severe restriction. A grasshopper (say), even one with human intelligence, could never tame a cat; the cat could never interpret the grasshopper's signals as those of a fellow creature.

So DREADCO technicians have devised a set of radio-controlled 'pet-loving robots' with various facial, ocular and vocal abilities. Each robot, guided by its unseen controller, attempts to feed, stroke, talk to and exchange friendly behaviour with an initially naive cat, dog or other animal. Those robots which succeed in making a pet of their animal will reveal the emotional signals recognized by that animal. In particular, they may prove a thesis Daedalus has held for some time. Our animal-taming abilities are severely limited by the immobility of our ears.

Cats, dogs, horses, rodents and many other creatures can swivel or shape their ears in different ways. We interpret these movements, instinctively and correctly, in emotional terms. Flattened or backwards-pointing ears suggest aggression or distress; upright and forward-pointing ones show interest and friendliness. Pet animals must be greatly discouraged by the absence of such signals from ourselves.

So Daedalus is also experimenting with a special skull-cap equipped with large electromagnetically adjustable 'ears'. A handset allows the wearer to switch them to whatever emotion he wishes to convey. With luck, the cap will enable its wearer to reach an emotional closeness with any animal that also uses ear signals.

DREADCO's ear-cap, together with any further prostheses indicated by the robot program, will transform our relations with ear-signalling animals. The most unruly dogs, aloof cats, treacherous goats and indifferent gerbils will at last respond to human affection. The acid test will be the utterly untameable Highland wild cat, reputed to hate everything and everybody on sight. If the cap can reduce this furious beast to a purring fireside moggy, a new lawn in the Garden of Eden will indeed have been opened.

David Jones

A *Velociraptor* wishbone

The 'wishbone' of birds comprises two clavicles fused into a structure¹ known as a furcula. In an influential 1926 book on bird origins by Heilmann², the furcula's supposed absence in dinosaurs was considered powerful evidence barring them from bird ancestry. A furcula has now been found in several theropod dinosaurs^{3,4}, but its absence in other theropods, and the uncertainty of whether this absence is real or an artefact of preservation, obscures the evolutionary history of this structure⁵. Here we report the discovery of a furcula in *Dromaeosauridae*, a group posited to be the closest relative of birds^{6,7}.

Among close bird relatives (non-avian Maniraptora), a furcula is known only in *Oviraptoridae*⁸. Although a furcula was tentatively reported to be present in an articulated skeleton of *Velociraptor mongoliensis* preserved embracing a *Protoceratops* skeleton^{3,6}, the only firsthand description of this specimen maintained that it is not present⁸, and our observations of that specimen confirm this. A furcula has not been identified in any of the other five described species of *Dromaeosauridae*. Other taxa either lack any evidence of clavicles or, in one troodontid specimen⁹, clavicles are present but not fused. But the rarity of articulated specimens, the similarity of the furcula to ribs and gastralia, and incomplete preservation conspire to obscure whether this absence is real.

A partial skeleton of *Velociraptor mongoliensis* (specimen IGM 100/976) was discovered at Tugrugeen Shireh, Mongolia, during the 1991 expedition of the Mongolian Academy of Sciences–American Museum of Natural History Expedition. The specimen consists of a nearly complete skull and the anterior part of the skeleton, preserving derived characters that allow definitive reference to *Velociraptor mongoliensis*¹⁰.

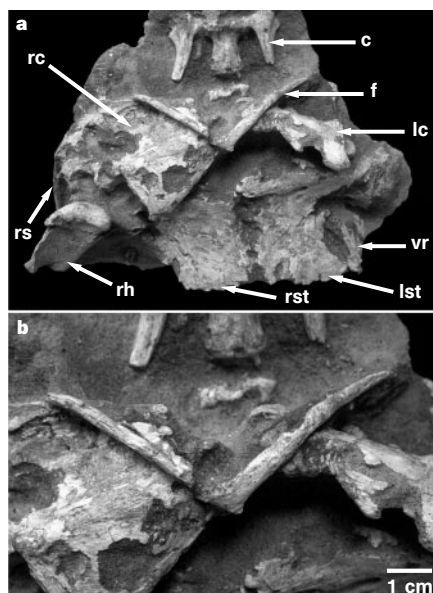


Figure 1 *V. mongoliensis* specimen IGM 100/976 with 'V'-shaped furcula. **a**, Ventral view of anterior trunk region. c, posterior cervical vertebrae; f, furcula; lc, left coracoid; vr, ventral ribs; lsc, left sternal plate; rst, right sternal plate; rh, right humerus; rs, right scapula; rc, right coracoid. **b**, Detail of same region.

The bones are heavily bioturbated, and many elements are punctured by borings similar to those often present in Djadokhta Formation sediments¹¹, but the specimen is otherwise very well preserved and articulated. The block that holds the furcula also preserves a pair of sternal plates, posterior cervical vertebrae, the proximal ends of both humeri, the coracoids and the right scapula. The specimen is large relative to others of this species, and the cervical ribs are nearly fused to the vertebral centra, indicating that the specimen is near full-size.

The furcula is in articulation between the scapulocoracoids anterior to the sternal

plates (Fig. 1), in the same position as in oviraptorids and birds. The bone is 'V'-shaped and very slender, much thinner than in oviraptorids and *Archaeopteryx*. In cross-section it is nearly circular. A short midline process, the hypocleidium, extends posterodorsally. The proximal, or epicleideal, process sweeps posterodorsally to articulate with the acromion of the scapulocoracoid. The proximal process of the furcula tapers to a point where it contacts the scapulocoracoid, and the posterior articulating surface contacting the acromion is smooth.

The broad distribution of a furcula among the non-flying relatives of birds^{4,5,8,12} indicates that its origin is not tied to the origin of flight. The furcula's use in powered avian flight is therefore a co-option of a structure already present before flight evolved in the lineage leading to modern birds⁴. The discovery of a furcula in a taxon previously known from an articulated specimen underscores the difficulty of assessing whether the apparent absence of this feature is real or an artefact of preservation or ontogenetic sampling⁵.

Mark A. Norell, Peter Makovicky
Department of Vertebrate Paleontology,
American Museum of Natural History,
79th Street at Central Park West,
New York, New York 10024-5192, USA

James M. Clark
Department of Biological Sciences,
George Washington University,
Washington DC 20052, USA

1. Baumel, J. J. & Witmer, L. M. *Publ. Nuttall Ornithol. Club* **23**, 45–132 (1993).
2. Heilmann, G. *The Origin of Birds* (Witherby, London, 1926).
3. Ostrom, J. *Biol. J. Linn. Soc.* **8**, 91–182 (1976).
4. Makovicky, P. J. & Currie, P. J. *J. Vert. Paleontol.* (in the press).
5. Bryant, H. & Russell, H. J. *Vert. Paleontol.* **13**, 171–184 (1993).
6. Gauthier, J. A. *Calif. Acad. Sci. Mem.* **8**, 1–55 (1986).
7. Holtz, T. J. *J. Paleontol.* **70**, 536–538 (1996).
8. Barsbold, R. *Sovmestnaya Sovetskoye-Mongol'skaya Paleontol. Ekspeditsiya, Trudy* **19**, 5–119 (1983).
9. Russell, D. A. & Dong, Z.-H. *Can. J. Earth Sci.* **30**, 2163 (1993).
10. Norell, M. A. et al. *J. Vert. Paleontol.* **12** Suppl. 45A (1992).
11. Jerzykiewicz, T. et al. *Can. J. Earth Sci.* **30**, 2180–2195 (1993).
12. Chure, D. J. & Madsen, J. H. *J. Vert. Paleontol.* **16**, 573–577 (1996).

Porous silica via colloidal crystallization

Microstructured porous silicas have potential applications in catalysis, separations, coatings, microelectronics and electro-optics, but methods for producing materials with uniform submicrometre pores have not been available. We have now developed a method in which modified colloidal crystals are used as templates for silica polymerization. This method yields products with highly uniform and structured pores of tuneable size in the submicrometre region.

Most previous syntheses of microstructured porous silicas have used supramolecu-

lar surfactant micelles as templates for silica polymerization, resulting in mesoporous M41S phases^{1–4}. This method is straightforward, but the pore cross-section is limited to about 10 nm or less by the size of the surfactant aggregates⁴. Recently, bacterial threads were mineralized to produce materials with oriented, elongated pores of micrometre size⁵. But the formation of inorganic materials with ordered uniform pores of arbitrary size and shape remains a challenge⁶.

We use crystalline structures assembled from colloid-sized particles ('colloidal crystals') as templates. Such crystals reveal basic principles of the interactions in the colloidal domain^{7,8}. New methods for controlling the formation^{9,10} and orientation¹¹ of the particle lattices are available, but the

crystals obtained have found little practical application, largely because the colloidal crystals are formed in liquid environments and either do not survive drying or have poor mechanical stability. Our scheme overcomes these problems by replicating the structure with a stable silica matrix.

Formation of colloidal crystals from monodisperse colloid particles is easily achieved by increasing their volume fraction in the vicinity of a flat wall¹². We filtered polystyrene latex microspheres (size range 200–1,000 nm), carrying either negative surface charges (sulphate) or positive charges (amidine), in suspension through a smooth narrow-pore membrane. When the particulate layers had been deposited, they could be washed and penetrated by different liquid

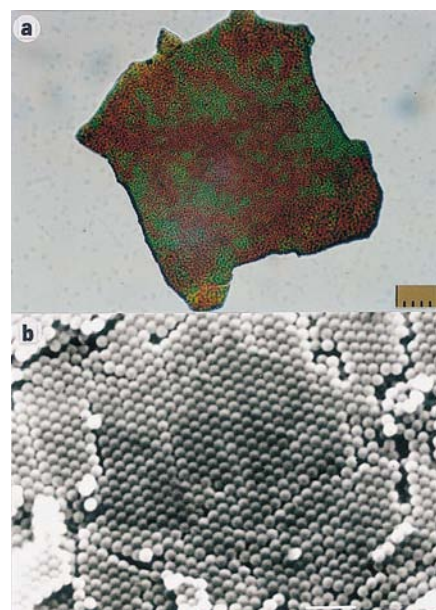


Figure 1 Micrographs of ordered latex templates. **a**, Optical microscopy in transmitted light. The different colours probably correspond to crystal domains of different symmetry or orientation. **b**, Scanning electron micrograph (SEM). Ordered latex layers display either hexagonal or cubic packing. Scale bars: 50 μm and 1 μm , respectively.

media. The latex particles accumulated slowly on the membrane surface, building up closely packed, ordered layers roughly 10 μm thick. The deposited crystalline layer could be broken and detached from the membrane surface for analysis by light and scanning electron microscopy (Fig. 1).

To induce silica polymerization the microsphere surfaces had to be functionalized *in situ* by adsorption of the surfactant hexadecyltrimethyl ammonium bromide (HTAB). We soaked the crystalline latex layers with 0.02 M HTAB solution for 20 min, then removed the excess unadsorbed surfactant by washing briefly with deionized water. We mineralized the cavities in the arrays by passing 0.5 M silica solution through the latex-covered filter. The permeability of the layers decreased as the polymerization process continued, so that flow through the filter stopped in less than one minute. When the silica solution had gelled inside the colloidal crystal layer, we removed the excess solution and dried the latex/silica composite under vacuum. The latex templates inside the polymerized silica were removed by heating at 450 $^{\circ}\text{C}$ for 4 h, leaving silica flakes of very low density as the final product.

Scanning electron microscopy shows that the material is built up of three-dimensional ordered arrays of uniform pores. Examples of the discrete morphology of the material are shown in Fig. 2. By varying the size of the latex microspheres used, we were able to produce organized materials with pore sizes ranging from about 150 nm to

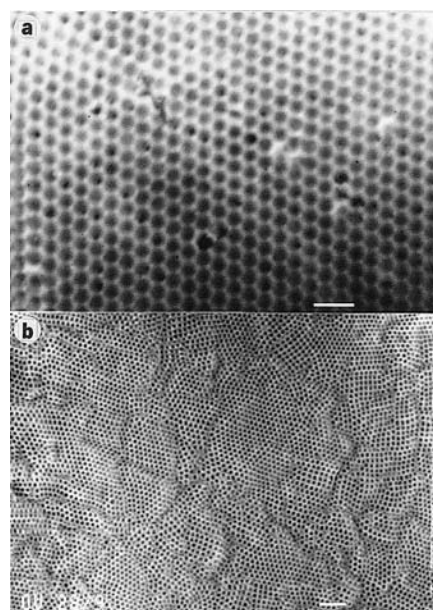


Figure 2 SEM of the microporous silica structures. Latex particles of **a**, 560 nm and **b**, 300 nm diameters were used as templates. Large ordered arrays of spherical cavities, representing a negative replica of the original colloidal crystal embedded in the silica, are seen. The silica flakes appear to be built up of many similar domains with different crystal orientations. Details of materials and methods are available on request from the authors. Scale bars, 1 μm .

1 μm . A comparison between the repeat units of the silica replicas and the original latex crystals showed that the baked materials had shrunk by 20–35%, a value that is higher than, but comparable to, that in the M41S mesoporous silicas¹.

Our results show that it is possible to obtain highly structured silica materials in which the pore size, shape and ordering can be precisely controlled within a wide range that has previously been unattainable. The method is powerful and controllable, and could be adapted for large-scale production.

**O. D. Velev, T. A. Jede
R. F. Lobo, A. M. Lenhoff**

Department of Chemical Engineering,
University of Delaware, Newark,
Delaware 19716, USA
e-mail: velev@che.udel.edu

1. Kresge, C. T., Leonowicz, M. E., Roth, W. J., Vartuli, J. C. & Beck, J. S. *Nature* **359**, 710–712 (1992).
2. Beck, J. S. *et al.* *J. Am. Chem. Soc.* **114**, 10834–10843 (1992).
3. Monnier, A. *et al.* *Science* **261**, 1299–1303 (1993).
4. Raman, K. N., Anderson, M. T. & Brinker, C. J. *Chem. Mater.* **8**, 1682–1701 (1996).
5. Davis, S. A., Burkett, S. L., Mendelson, N. H. & Mann, S. *Nature* **385**, 420–423 (1997).
6. Mann, S. & Ozin, G. A. *Nature* **382**, 313–318 (1996).
7. Arora, A. K. & Rajagopalan, R. in *Ordering and Phase Transitions in Charged Colloids* (eds Arora, A. K. & Tata, B. V. R.) 1–13 (VCH, New York, 1996).
8. Larsen, A. E. & Grier, D. G. *Nature* **385**, 230–233 (1997).
9. Trau, M., Saville, D. A. & Aksay, I. A. *Science* **272**, 706–709 (1996).
10. Denkov, N. D. *et al.* *Nature* **361**, 26 (1993).
11. van Blaaderen, A., Ruel, R. & Wiltzius, P. *Nature* **385**, 321–324 (1997).
12. Davis, K. E., Russel, W. B. & Glantschnig, W. J. *Science* **245**, 507–510 (1989).

The same prion strain causes vCJD and BSE

Epidemiological and clinicopathological studies, allied with pathological prion protein (PrP^{Sc}) analysis, strongly support the hypothesis that the human prion disease new variant Creutzfeldt–Jakob disease (vCJD) is causally related to bovine spongiform encephalopathy (BSE)^{1,2}, but considerable controversy remains. Distinct prion strains are distinguished by their biological properties on transmission to laboratory animals and by physical and chemical differences in PrP^{Sc} strains. We now find that the biological and molecular transmission characteristics of vCJD are consistent with it being the human counterpart of BSE.

We studied transgenic mice expressing only human PrP (HuPrP^{+/+} *Prn-p^{0/0}*), which have been shown to lack a species barrier to human prions from one iatrogenic CJD case³, comparing them with non-transgenic (FVB) mice. All of 16 further CJD cases, encompassing a wide range of clinicopathological phenotypes, all three PrP^{Sc} types reported in sporadic and acquired prion diseases² and all *PRNP* genotypes at polymorphic codon 129, a key determinant of genetic susceptibility to human prion diseases^{4–6}, were transmitted to these transgenic mice.

Almost all inoculated transgenic mice contracted disease with similar short incubation periods, consistent with a lack of species barrier to these isolates (Table 1). These transgenic mice express human PrP homozygous for valine at codon 129. However, there was no significant difference in mean incubation periods between inocula of the different codon 129 genotypes. PrP^{Sc} typing of these transmissions showed that the same prion types seen in sporadic and iatrogenic CJD (types 1–3) are produced, distinct from that seen in vCJD (type 4)². Only occasional transmissions, at longer and variable incubation periods, were seen in FVB mice.

In contrast, efficient transmission of vCJD to FVB mice was observed (Table 1) although incubation periods were prolonged. Conversely, the attack rate of vCJD in the transgenic mice was reduced in comparison to typical CJD, and incubation periods were generally more variable and prolonged. Mean incubation periods to these six vCJD cases were similar in both types of mice. The clinical course in vCJD-inoculated transgenic mice was much longer than in transmissions of typical CJD. vCJD in humans is also associated with a long clinical duration¹. Some mice, as well as showing typical neurological features, persistently walked backwards. This unusual clinical sign was not seen in transmissions of typical CJD, fatal familial insomnia or other

inherited prion diseases⁷.

BSE transmits efficiently to FVB mice³, albeit with prolonged and variable incubation periods (Table 1) which fall to a consistent short incubation period of around 140 days on second passage (data not shown). Transmissions of BSE into the transgenic mice did not occur at incubation periods well beyond those of classical CJD³, but we have now observed transmission with much longer incubation periods (Table 1). These transmissions resembled those of vCJD with a long clinical duration and backwards walking in some animals as well as the otherwise typical clinical features of mouse scrapie.

There were striking similarities in PrP deposition patterns between BSE- and vCJD-inoculated animals (detailed neuropathological studies will be published elsewhere). Such patterns are determined by host genotype as well as by agent strain⁸. We saw distinct patterns in the two types of mice, but, in each case, vCJD and BSE produced closely similar patterns. In vCJD- and BSE-inoculated non-transgenic mice, there were PrP plaques and diffuse PrP deposition. In vCJD- and BSE-inoculated HuPrP^{+/+} *Prn-p⁰* transgenic mice we saw a predominantly pericellular pattern of PrP immunostaining (data not shown). PrP plaques are a rare feature of prion disease in mice. Occasional mock-inoculated transgenic mice showed weaker and less extensive pericellular PrP immunostaining, probably reflecting the high level of PrP^C overexpression in these mice. Western blotting for PrP^{Sc} was negative in all these controls.

We performed western blot analysis to determine the PrP^{Sc} types produced in these transmissions. We have previously shown that the PrP^{Sc} type seen in vCJD (type 4) has a ratio of glycoforms closely similar to that of BSE passaged in several other species². vCJD-inoculated FVB mice produced mouse PrP^{Sc} with type 4-like glycoform ratios and fragment sizes indistinguishable from those in BSE-inoculated FVB mice (Fig. 1a,b).

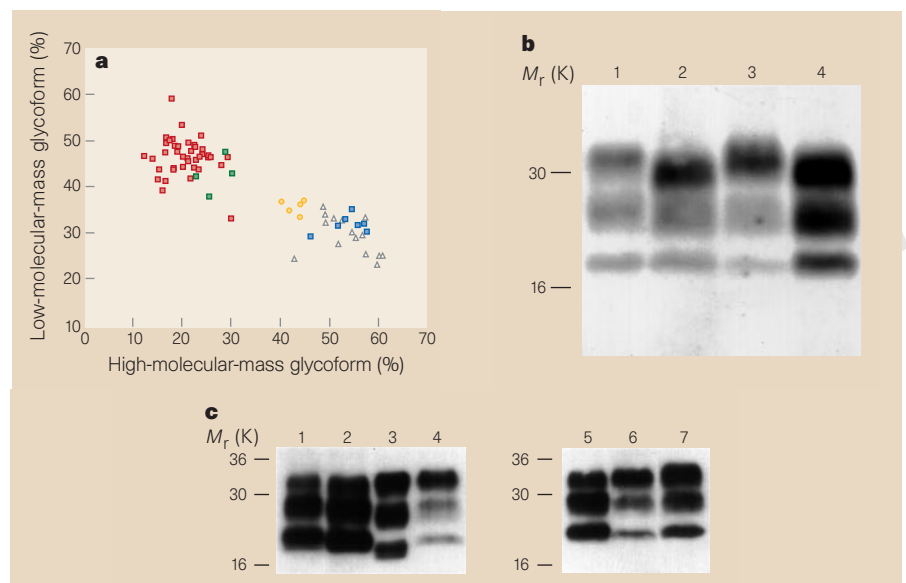


Figure 1 Transmission of prion diseases to mice. **a**, Scatter graph of proportions of protease-resistant PrP in the high-molecular-mass (di-glycosylated) and low-molecular-mass (mono-glycosylated) glycoforms in individual human cases and FVB mice with experimentally transmitted CJD, vCJD or BSE. Sporadic and iatrogenic CJD cases (PrP^{Sc} types 1–3), red squares; vCJD, yellow circles; transmissions of typical CJD to FVB mice, green squares; BSE to FVB mice, blue squares. Transmissions of vCJD to FVB mice, open triangles. **b,c**, Western blots of brain homogenates after pre-treatment with proteinase K using anti-PrP polyclonal antibody 95-108 (ref. 15) (**b**) or anti-PrP monoclonal antibody 3F4 (**c**). Methods were as in ref. 2 except that for PrP glycoform analysis a chemifluorescent substrate (ECF, Amersham) was used and ratios analysed on a Storm 840 Phosphorimager (Molecular Dynamics). **b**, Transmission of vCJD and BSE to non-transgenic (FVB) mice. Lane 1, human vCJD; 2, vCJD-inoculated FVB mouse (same case as lane 1); 3, BSE; 4, BSE-inoculated FVB mouse (same case as in lane 3). **c**, Transmission of vCJD to HuPrP^{+/+} *Prn-p⁰* transgenic mice. Lane 1, human CJD, type-2 PrP^{Sc}; 2, transgenic mouse inoculated with CJD case from lane 1 showing type-2 pattern; 3, human vCJD case, type-4 PrP^{Sc}; 4, transgenic mouse inoculated with vCJD from lane 3 showing type-5 pattern; 5, human CJD case, type-2 PrP^{Sc}; 6 and 7, type-5 PrP^{Sc} pattern in vCJD-inoculated transgenic mice.

In transmission of vCJD to HuPrP^{+/+} *Prn-p⁰* transgenic mice, where human PrP^{Sc} is generated, fragment sizes in inoculum and host can be directly compared. Again the PrP^{Sc} produced had type 4-like glycoform ratios. However, the fragment sizes differ from those in the inoculum and were indistinguishable from those in the type-2 PrP^{Sc} pattern² (Fig. 1c). We have designated this new pattern type 5.

A change of fragment size on passage in

mice of a different codon 129 PrP genotype than the inoculum has been reported previously². Type-1 PrP^{Sc}, seen in CJD cases of 129MM *PRNP* genotype, consistently converts to type-2 PrP^{Sc} on passage in these transgenic mice expressing 129VV human PrP. The glycoform ratios of the original inoculum are also maintained². Abrupt changes in the biological properties ('mutation') of murine scrapie strains on passage in mice of different genotypes are well recognized⁹. We have not, however, been able to show PrP^{Sc} by western blotting in BSE-inoculated HuPrP^{+/+} *Prn-p⁰* transgenic mice. This may reflect culling of many of these mice soon after clinical diagnosis rather than at a more advanced clinical stage. Though transmission of prion diseases without detectable PrP^{Sc} on primary passage has been reported^{7,10}, it will be important to confirm transmission by second passage studies.

The prion titres in these primary inocula are unknown but may be higher in the human cases, because cattle with BSE will have been culled before the terminal stages of disease. However, on clinical, pathological and molecular criteria, vCJD shows remarkable similarity in its transmission characteristics to BSE, and is quite distinct from all other forms of sporadic and acquired CJD. These data provide compelling evidence that

Table 1 Incubation periods for transmission of prion diseases to transgenic and wild-type mice						
Type of inoculum	PRNP codon 129	No. of inocula	Affected/ inoculated	Incubation period (days)	Affected/ inoculated	Incubation period (days)
spCJD	MM	9	66/67	210 ± 4	0/60	>450 or >600
spCJD	VV	1	5/5	337 ± 11	1/5	471
spCJD	MV	2	15/15	218 ± 2	1/9	257
iCJD (GH)	MM	1	7/7	211 ± 5	1/5	318
iCJD (GH)	MV	1	4/4	195 ± 9	1/5	367
iCJD (GH)	VV	1	5/5	193 ± 4	0/5	>600
iCJD (DM)	MM	1	4/4	204 ± 6	2/4	569, 569
iCJD (G)	VV	1	8/8	187 ± 4	0/5	>600
vCJD	MM	6	25/56	228 ± 15	33/43	371 ± 17
BSE		5	10/26	602 ± 50	21/24	466 ± 26

Incubation periods (mean±s.e.m.) in HuPrP^{+/+} *Prn-p⁰* transgenic mice and non-transgenic (FVB) mice. Methods and PRNP analysis were as ref. 2. All CJD cases were neuropathologically confirmed. BSE inocula were pooled brainstem and four individual confirmed cases. Most mice were examined by neuropathology and/or western blotting. Incubation periods are not given for mice that died without definite clinical diagnosis, despite neuropathological or western blot evidence. Numbers affected/inoculated exclude mice dying after inoculation or from intercurrent illness. Controls inoculated with PBS alone did not develop prion disease. spCJD, sporadic CJD; iCJD, iatrogenic CJD; GH, growth hormone; DM, dura mater; G, gonadotrophin; MM, methionine homozygote; MV, methionine/valine heterozygote; VV, valine homozygote.

BSE and vCJD are caused by the same prion strain. Taken together with the temporal and spatial association of vCJD with BSE but not with scrapie or other animal prion diseases, and BSE transmission studies in macaques¹¹, this strongly suggests that vCJD is caused by BSE exposure. The theoretical possibility that both BSE and vCJD arise from exposure to a common unidentified source appears remote.

The production of a distinct molecular strain type on transmission of vCJD to mice expressing valine 129 human PrP suggests that BSE transmitted to humans of this genotype might produce a similar strain. Such cases may differ in their clinical and pathological phenotype to vCJD, but could be identified by PrP^{Sc} typing.

Although it has been argued that the species barrier resides in PrP primary structure differences between donor and host¹², our data emphasize that strain type can be as important. As prion propagation involves interactions between PrP^{Sc} and host PrP^C, and strains are associated with differences in PrP conformation and glycosylation^{2,13}, such PrP interactions may be most efficient if the interacting proteins are not only of the same sequence but have similar conformational preferences and glycosylation. Mismatch of codon 129 between inoculum and HuPrP^{+/+} Prn-p^{0/0} mice does not significantly affect CJD transmission, but this could differ for BSE. All vCJD cases have been 129MM genotype (ref. 14 and unpublished data). Although our 129VV mice are much less susceptible to BSE than to typical CJD, suggesting a substantial species barrier, 129MM human PrP mice could be more susceptible.

Andrew F. Hill, Melanie Desbruslais

Susan Joiner, Katie C. L. Sidle

Ian Gowland, John Collinge*†

Prion Disease Group, Neurogenetics Unit,
Imperial College School of Medicine at St Mary's,
London W2 1PG, UK

*and Department of Neurology, St Mary's Hospital,
London, W2 1NY, UK

Lawrence J. Doezy, Peter Lantos

Department of Neuropathology,
Institute of Psychiatry, London SE5 8AF, UK

†To whom correspondence should be addressed at the Prion
Disease Group.

- Will, R. G. *et al. Lancet* **347**, 921–925 (1996).
- Collinge, J., Sidle, K. C. L., Meads, J., Ironside, J. & Hill, A. F. *Nature* **383**, 685–690 (1996).
- Collinge, J. *et al. Nature* **378**, 779–783 (1995); addendum **389**, 526 (1997).
- Baker, H. F. *et al. Lancet* **337**, 1286 (1991).
- Collinge, J., Palmer, M. S. & Dryden, A. J. *Lancet* **337**, 1441–1442 (1991).
- Palmer, M. S., Dryden, A. J., Hughes, J. T. & Collinge, J. *Nature* **352**, 340–342 (1991).
- Collinge, J. *et al. Lancet* **346**, 569–570 (1995).
- Bruce, M. E. *et al. J. Gen. Virol.* **72**, 595–603 (1991).
- Bruce, M. E. *Br. Med. Bull.* **49**, 822–838 (1993).
- Lasmézas, C. I. *et al. Science* **275**, 402–405 (1997).
- Lasmézas, C. I. *et al. Nature* **381**, 743–744 (1996).
- Prusiner, S. B. *et al. Cell* **63**, 673–686 (1990).
- Telling, G. C. *et al. Science* **274**, 2079–2082 (1996).
- Collinge, J. *et al. Lancet* **348**, 56 (1996).
- Piccardo, P. *et al. J. Neuropathol. Exp. Neurol.* **56**, 589 (1997).

'Male-stuffing' in wasp societies

Intracolony aggression within and between castes of social insects is common^{1–3}. We have observed an unusual aggressive interaction between nestmates of the paper wasp *Polistes dominulus*. In response to foragers returning to the colony, females (workers) initiate aggressive encounters with males culminating with the male being forced head-first into an empty nest-cell ('male-stuffing'). 'Stuffed' males are unable to feed, so the behaviour seems to ensure that food is preferentially channelled to larvae, which are likely to be more closely related to the workers than are the adult males.

We observed two categories of stuffing. 'Initial stuffing' (Fig. 1) began with antenna-to-antenna contact and was followed by grappling, biting, and sting-threats. The aggressor then forced the recipient head-first into an empty cell. 'Repeated stuffing' was characterized by biting and pushing the abdomen of an individual whose head and thorax were already inside a cell.

We studied the behaviour by transcribing and analysing 26 hours of videotape. We saw stuffing behaviour only in colonies containing males ($n=5$ colonies) and not in those without ($n=6$ colonies; sexed by antennal morphology²; $\chi^2=21$, $P<0.001$). Stuffing was directed exclusively at males, despite their being greatly outnumbered by females (1:4.21) in colonies of both sexes (binomial test, $P<0.0001$). Of 66 stuffing events, 46 were directed at males from that colony (identified by marking them at eclosion); the remainder were of unknown origin. Queens ($n=5$) did not stuff males (0/66 events; binomial test, $P<0.1$). All stuffing was done by workers other than the returning forager.

Initial stuffing occurred soon after the return of a forager, whereas repeated stuff-

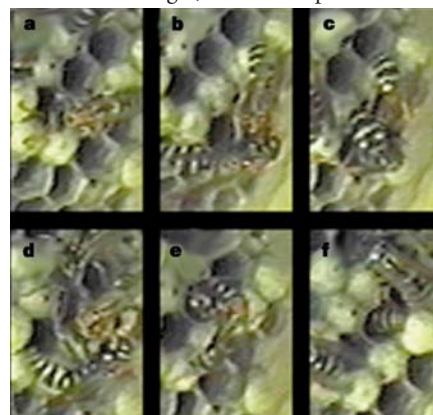


Figure 1 Initial male-stuffing. **a**, Male on the comb. **b**, Female (worker) approaches and antennates him, **c**, followed by biting and sting-threats. **d**, She stuffs him into an empty cell, **e**, and pushes on his abdomen. **f**, Male in the cell.

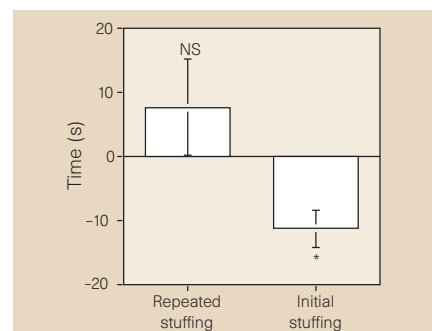


Figure 2 Difference between the time from most recent male arrival until stuffing and half the average interval between returns. A value of 0 is expected if male-stuffing is random with respect to arrivals. Initial stuffing ($n=32$) occurred shortly after a nestmate returned ($\bar{t}=18.86 \pm 2.89$ s; Wilcoxon signed-rank test: $Z=-3.20$, $*P<0.01$), but repeated stuffing ($n=34$) occurred randomly with respect to arrivals ($\bar{t}=39.88 \pm 7.51$ s; $Z=-0.18$, $P>0.8$; NS). Means $\pm$ s.e.m.

ing occurred at random times (Fig. 2). Males that had been repeatedly stuffed remained in cells 6.35 times longer ($\bar{t}=384.29 \pm 43.01$ s; mean time $\pm$ s.e.m.) than the mean time between forager arrivals ($\bar{t}=60.53 \pm 2.25$ s; $n=833$). Thus, stuffing may function to preclude males from gaining access to resources gathered by the workers.

Limiting food consumption by males may maximize the inclusive fitness of workers, who should direct their help towards closely related kin^{4,5}. Feeding future reproductive females provides a larger fitness pay-off than feeding adult males⁶. Workers from a colony containing one singly mated queen have a relatedness to sisters of 0.75. Workers are only related by 0.25 to brothers, 0.375 to nephews (worker-produced males) and are unrelated to immigrant males.

Assuming that female larvae are present, workers are more closely related to reproductive-destined larvae than to adult males. Even in circumstances where workers are, on average, equally related to male and female nestmates (such as brothers and half-sisters when the queen has mated more than once), feeding needy larvae may provide a larger inclusive fitness pay-off than feeding adult males, which can forage for themselves. Preferential channelling of resources to larvae, by stuffing males, may maximize the genetic self-interest of worker wasps.

Philip T. Starks, Emily S. Poe

Section of Neurobiology and Behavior,
W311 Seeley G. Mudd Hall, Cornell University,
Ithaca, New York 14853-2702, USA
e-mail: pts3@cornell.edu

- Ratnieks, F. L. W. Thesis, Cornell Univ., Ithaca (1989).
- Reeve, H. K. in *The Social Biology of Wasps* (eds Ross, K. G. & Matthews, R. H.) 99–148 (Cornell Univ. Press., Ithaca., 1991).
- Gamboja, G. J. in *Natural History and Evolution of Paper-Wasps* (eds Turillazzi, S. & West-Eberhard, M. J.) 161–177 (Oxford Univ. Press., 1996).
- Hamilton, W. D. *J. Theor. Biol.* **7**, 1–16 (1964).
- Hamilton, W. D. *J. Theor. Biol.* **7**, 17–52 (1964).
- Hamilton, W. D. *Annu. Rev. Ecol. Systemat.* **3**, 193–232 (1972).

Return of the repressed

Scientific Knowledge: A Sociological Analysis

by Barry Barnes, David Bloor and John Henry

University of Chicago Press/Athlone: 1996.
Pp. 230. \$38, £42 (hbk); \$15.95, £16.95 (pbk)

Science

by Steve Fuller

University of Minnesota Press/Open University Press: 1997. Pp. 176. \$37.95, £35 (hbk); \$14.95, £9.99 (pbk). To be published on 31 October

Ziauddin Sardar

Scepticism has a long history of dethroning disciplines that try to monopolize truth. It knocked religion from its pedestal, and then went on to undo philosophy. Now its chief target is science. These two books are the latest in a long line of onslaughts on the claimed integrity and truth of science.

Received wisdom sees science as neutral and value-free and in no need of sociological analysis. Scientific knowledge is precisely that which is not social; it is universal and believed because it is true and rational.

David Bloor and Barry Barnes were among the first to challenge this view. In his seminal work *Knowledge and Social Imagery* (1976), Bloor argued that sociology had a fundamental role in analysing how 'true' scientific knowledge is produced. Bloor and Barnes, both associated with the 'Edinburgh school', were the chief architects of what came to be known as the 'strong programme' in the sociology of scientific knowledge, or SSK (see Briefing in *Nature* 387, 331–335; 1997).

It had four basic elements. First, the Edinburgh school argued, the goal of SSK is to discover the conditions that bring about states of knowledge; these conditions can be economic, political and social as well as psychological. Second, SSK had to be impartial in its selection of what is studied, giving equal emphasis to true and false knowledge, the successes and failures of science, and rational and irrational enquiries. Third, there should be consistency or 'symmetry' in the explanations sought for selected instances of scientific knowledge; one could not use, for example, a sociological cause to explain a 'false' belief and a rationalist cause to explain a 'true' belief. Fourth, the models of explanation of SSK had to be applicable to sociology itself.

When first formulated, SSK and the strong programme were truly radical; indeed they were seen as subverting science. They generated, in equal measure, controversy about and advances in our understanding of the nature of science. *Scientific Knowledge* extends and updates this understanding while trying to show how SSK is actually practised.

Barnes and his coauthors argue that it is not so much the observations in science that are 'theory-laden' (as is claimed by many proponents of social studies of science) but rather the reports of the observations. Unlike the more radical critics, they accept that perception itself can be "stable and autonomous".

They show, however, that the way in which an observation is reported depends on the tradition in which a scientist works. So two scientists working in different (scientific) traditions may observe the same results but report and interpret them in different ways. This argument is clarified and illustrated by a detailed examination of the famous and controversial oil-drop experiment of the US physicist R. A. Millikan and the challenge to his work by the Viennese physicist Felix Ehrenhaft.

Moreover, the authors argue, theories themselves are not fixed in time, nor can they be identified with a set of fixed statements. Association of theories with the names of great scientists—Newton's theory of gravitation, Einstein's theory of relativity—creates this illusion. It is better to think of scientific theories as evolving institutions: a detailed examination of Mendel's theory of heredity shows us how many twists and turns it has taken since Mendel first formulated it.

The Edinburgh school has always argued that 'experience' and 'reality' are actually 'out there'. Realism should not be opposed but illuminated by sociological enquiry. The new conceptual tool they offer for elucidating realism is "sociological finitism", which is a way of looking at how 'words' are connected to the 'world'.

Finitism suggests that all scientific terms and concepts are open-ended—"no specification or template or algorithm fully formed in the present" is "capable of fixing the future correct use of the term"—and emphasizes the conventional character and sociological interest of scientific classifications. Finitism also suggests that the boundaries between scientific disciplines, as well as the demarcation between what science is and what it is not, are liable to change with changing situations. In this way it is conceivable that a shift in our conception of what science is may lead to incorporation of what is currently dismissed as nonscience into science: astrology, acupuncture, parapsychology and so on.

This is the kind of sociology of knowledge that most rank-and-file scientists would grudgingly endorse. The authors of *Scientific Knowledge* do not seem to acknowledge any 'social interest' in science outside science itself—there is no sense of larger social forces operating on science beyond those the scientists witness.

Although they continue the project of demystification of science that became prominent with Thomas Kuhn, they see SSK as a "part of the project of science itself" and wish to use the methodology of science to analyse science. In other words, the authors are presenting themselves as a Kuhnian 'paradigm' and claiming to do normal science.

By accepting the existence of an unproblematic reality that is successfully explored in science, the Edinburgh school is, despite its radical approach, still fundamentally with the positivist camp. But Steve Fuller is much closer to full-blown scepticism.

In *Science*, he accuses the scientific community of a basic ignorance, saying that it has "misunderstood something significant about the social conditions that enabled its continued existence". By contrast, the public, whose 'understanding of science' is officially believed to need improvement, is accepted by Fuller as having a fundamentally sound perception of science. The campaign for 'public understanding of science', he reasonably asserts, is "our latest moral panic".

He rejects the positivist interpretation of science as well as seeking to undermine Kuhn's image as some sort of revolutionary. Kuhn's "little democracy" picture of science emerges from the view that once scientists agree on a particular paradigm as the basis for further enquiry, it is presented to the public as "self-contained expertise" requiring deference and obedience. The entire exercise serves a double purpose: scientists "retain their autonomy by displaying their authority".

But Fuller lives in a world where questions of history, power and Eurocentrism are very real. Barnes and his coauthors seem to be unaware of feminism and assessments of science from non-Western civilization; and as far as they are concerned, the Marxist assessment of science has never existed. Fuller sees science as a product of a particular civilization and therefore draws his material from a



SIDNEY HARRIS

broad cultural and historical range. He is just as interested in looking at science from the inside as from the outside: from the perspective of Islam, China and Japan.

Fuller even looks at science from the perspective of visiting Martians who are well-versed anthropologists. They start by wondering whether they should disguise themselves as "a member of a group that is essential to reproducing the social order of science but traditionally marginal to its power structure" such as "women, ethnic minorities, postdoctoral fellows and technical support staff". Eventually they decide to use conventional categories of religious thought to understand science.

Their conclusions: science does a good job of reminding scientists what it is like to be perfect but provides little guidance on how to improve; just because humans appeal to science to justify their beliefs does not mean that these beliefs are rational; scientists use "methodological ventriloquism" to project their descriptive and explanatory categories on an indifferent reality; and democratic governments use the appeal of scientific authority to control and coerce people.

To drive his points further, Fuller adds a couple of appendices to this chapter. The first inverts the famous norms of good scientific practice — universalism, communism, disinterestedness and organized scepticism — suggested by Robert Merton, the well known theorist of science. Fuller's Martians see the norms of science in terms of cultural imperialism, mafiosism, opportunism and collective irresponsibility. The second appendix deconstructs the *Science Citation Index* and reveals its subjective and "hyper-real" character and its manipulation by "epistemic cartels" and "strong chains of interpretive commands".

Science is a swashbuckling book. Fuller's formidable scholarship takes no prisoners; his project is not just to demonstrate the end of scientific knowledge but also to speed up its demise. Although the future of Western science is doubtful, however, traditional sciences in non-Western civilizations may come to the fore.

The emphasis on morality and ethics in science in Islam and the notions of mystery, duality and incoherence in Chinese science are signposts for the future. The Japanese have already demonstrated that the non-West does not need to retrace the history of science in the West or accept and adopt its

philosophical and metaphysical principles to evolve a particularly local scientific culture. What now seems 'reactionary' and rather untenable, from the Western perspective, could usher in the next scientific revolution as 'the return of the repressed'.

The discussion of science in non-Western civilisation is perhaps the weakest part of Fuller's book. For example, in examining science in Islam, he chooses to concentrate on the Islamic revivalists of the early and mid-twentieth century. As their perception of science was thoroughly coloured by their experience of colonialism, and because, despite their revivalist rhetoric, they were in awe of Europe's scientific success, the discussion is somewhat skewed. Moreover, given the truly monumental critiques of science developed in India, the total absence of Indian science in the book is unfortunate. Fuller will find kindred spirits in such giants of Indian sociology of knowledge as Jit Singh Uberio, Ashis Nandy, Claude Alvares, Shiv Viswanathan and Veena Das.

Both books claim to guide the reader through the minefield of the sociology of knowledge. However, whereas Fuller provides a wide-ranging tour of the whole gamut of the social studies of science, Barnes and his coauthors appear frozen in history.

Scientific Knowledge appears as a triumphalist account of SSK as it was 20 years ago, presented as the core of science studies. One cannot discover from *Scientific Knowledge* just how much science studies have developed and how huge and interdisciplinary an enterprise the subject has become. Moreover, the social sciences themselves have been radically transformed in the past few decades. They have been cured of the 'physics envy' that Barnes and his coauthors still rely on.

The Edinburgh school has been rightly criticized for professing a 'voodoo' sense of social causality of science that relies on standards of necessary and sufficient conditions so stringent that not even physics could satisfy them. But it is this acceptance of the physics model of explanation and causation that will appeal to many scientists and philosophers when they argue against social studies of studies.

By contrast, Fuller demonstrates the true importance of the sociological, philosophical and historical analyses of science by driving home the implications for the role of science in contemporary society. As scepticism comes home to roost, scientists need to read both books to gain a full appreciation of the enemy's terrain. Yesterday's sociological radicals have become today's ideological conservatives. It is an indication that the revolution is far from over. □

Ziauddin Sardar, visiting professor of science and technology policy at Middlesex University and consulting editor of *Futures*, is at 1 Orchard Gate, London NW9 6HU, UK.

Stormtrooper in stilettos

Figments of Reality: The Evolution of the Curious Mind

by Ian Stewart and Jack Cohen
Cambridge University Press: 1997. Pp. 344.
£16.95, \$24.95

Henry Gee

The lives of many impressionable graduate students (myself included) were enriched by Douglas Hofstadter's baroque confection *Gödel, Escher, Bach: The Eternal Golden Braid*, a self-styled fugue on the subject of mathematics, intelligence and consciousness. Hofstadter argues (I think) that any logical system that can encompass the human mind will, of necessity, contain undecidable propositions. That is why we will always be able to outsmart a computer.

But what I remember most about the book is its interludes — dialogues between Achilles and the tortoise of Zeno's paradox, placed to illuminate the text. The pair are occasionally joined by other characters such as Aunt Hilary, the emergent group-mind of an anthill, whose thoughts are encoded by the seemingly aimless movements of individual ants.

Figments of Reality owes a large (and acknowledged) debt to *Gödel, Escher, Bach*. (It also contains a lot about ants.) The fruit of a continuing complicit collaboration between the mathematician Ian Stewart and the reproductive biologist Jack Cohen, both also avowed science fiction buffs, it aims to show that human consciousness is as evanescent as any of Aunt Hilary's musings. The mind, they say, results from the complicit continuing interaction between human beings and culture.

The word 'complicit' is deliberate. More than a neologism, it is a central theme of *Figments*. Complicity is what happens when two systems interact to produce a third system, often superficially simple, whose behaviour tends to change the two interacting systems in a continuing process of feedback — the results of which cannot be predicted from a simple sum of the first two systems. The complicit Stewart and Cohen first set this out in their earlier and less elegant *Collapse of Chaos*. In many ways, *Figments* is the stiletto to *Collapse's* blunt instrument.

Complicity stands in contrast to reductionism, the notion that, ultimately, everything can be explained by an equation printable on a T-shirt. Moving upwards from the 'theory of everything' rapidly leads to an unresolvable confusion of specifics. Conversely, moving down from the 'real world' in an effort to find universal generalizations leads to a fragmentation into individual subsystems.

As scepticism comes home to roost, scientists need to read both books to gain a full appreciation of the enemy's terrain.

The two never meet — in between lies what Stewart and Cohen call “ant country”, the region where the complicit interaction of systems produces emergent phenomena that cannot be predicted by reductionist methods. The ‘ant’ business comes from a cellular automaton called Langton’s ant, which has inexplicably complex behaviour, entirely unpredictable from the moronically simple rules by which it operates.

Pomposity is the peril likely to afflict any book of this sort. The casual chat of Gödel, Escher, Bach goes too far the other way. *Figments*, by contrast, is alarmingly, almost full-frontally, direct. Stewart and Cohen assault (that *is* the right word) all the big questions (the origin of life, the nature of progression in evolution, the mind–body problem, the origin of consciousness and so on) with gleeful expedition. For example, their debunking of Roger Penrose’s idea of quantum consciousness as hopelessly reductionist is the literary equivalent of being taken round the back of the pub and mugged.

Figments even has its own version of Achilles and the tortoise. These are the Zarathustrans, amusing bird-like aliens who operate through an eightfold group-mind. “Octimality” is the watchword of the Zarathustrans: they even have their own theory of everything that fits on a Z-shirt (the equation is $E=8$). Stewart and Cohen’s side-long looks at life, the Universe and everything are paralleled by interludes in which the eight crew-members of the Zarathustran survey vessel *Watcher of Moons*, orbiting Earth, consider the same points — and reach intriguingly different conclusions. Amusement turns to terror as the Zarathustrans consider, seriously, whether humanity should be destroyed on the grounds of unoctimality.

At every step, *Figments* questions our placidly received wisdom. Its metaphors often come from science fiction. It is sad that many scientists dismiss science fiction as juvenilia. Ninety per cent of sci-fi is daft, naturally — but so is 90 per cent of everything in the end. And the best sci-fi raises the most scientific questions in readers’ minds, questions that always start “What if...?”

Like all books, *Figments* has its faults, mainly that the authors tend to get carried away by their own metaphors — something that leads them into a view of evolution that looks suspiciously orthogenetic. Again, if life on Earth is monophyletic, using morphological analogues as a way of predicting alien morphology is suspect. Finally, the word ‘oops’ has no place outside cartoon speech balloons. Such niggles are unoctimal, as ultimately (and fittingly) *Figments* is a great deal more than the sum of its parts.

Figments will not appeal to everyone. It will hold few charms for anyone homozygous for the *humourless* mutagene, or without much imagination. But science without humour and imagination seems rather point-

less — just like books without pictures or conversation were to Alice. We need imagination to frame hypotheses, to wonder what is round the next corner, to ask “What if...?”. And anyone without a sense of humour should be a merchant banker, not a scientist — the pay is better, for a start.

For the rest of us, *Figments* is frighteningly readable, and a worthy successor to Gödel, Escher, Bach. □

Henry Gee is chief science writer on Nature.

A practical approach to bodybuilding

Principles of Tissue Engineering

edited by Robert P. Lanza, Robert Langer and William L. Chick

Academic: 1997. Pp. 808. \$125, £90

Nicholas A. Peppas

The design and engineering of new tissue for functional organs has been a fairly recent endeavour of scientists and engineers working on cell growth and proliferation. The term ‘tissue engineering’ has been around for more than a decade. In 1987, the US National Science Foundation offered the first request for proposals on tissue engineering, and the resulting programme set up not much later supported the research of some of the early pioneers in the field, including one of the coeditors of this volume, Robert Langer.

The premise of tissue engineering is simple: to replace diseased organs with newly grown, functional organs rather than with structured biopolymers and other biomaterials. Naturally, the characteristics of cell



A bag of silicon implanted in a female chest wall, defying the laws of gravity. From *The Secret Family: Twenty-Four Hours Inside the Mysterious World of Our Minds and Bodies* by David Bodanis. Simon & Schuster, \$27.50 (hbk).

growth and differentiation become important in designing a functional organ. Although *in vitro* control of tissue development is necessary for selection of appropriate cells and scaffolds, *in vivo* synthesis is, of course, the heart of the field.

This new treatise on the principles of tissue engineering is essential for anyone working in the field. It is a vast, detailed and beautifully presented analysis of the cellular principles, *in vitro* and *in vivo* behaviour, modelling and applications of tissue engineering. The subject is broader than its immediate application, because successful design and engineering of organs requires an appreciation of peripheral sciences such as biology, materials science, and chemical and mechanical engineering. The book provides sufficient background in all these fields for a reasonably trained scientist to appreciate any problem that may arise in tissue engineering.

The coverage of the subjects is detailed and clearly annotated, with emphasis on the basics of cell growth and differentiation, *in vitro* control of tissue development, *in vivo* synthesis of tissues, the use of biomaterials as scaffolds in tissue engineering, transplantation issues and applications in the cardiovascular system, the gastrointestinal system, the kidney, reconstruction of cornea and pancreas, growth of cartilage and bones, and nervous tissue regeneration as well as dental and skin applications.

With a book this size, one cannot but be impressed by the coverage. Regrettably, the multi-authored approach (88 authors contributing 48 chapters) creates some problems despite the editors’ considerable efforts. Stricter editing would have avoided occasional awkwardness, as in oversimplifications about lymphocyte engineering (p. 530), self-congratulatory comments by authors (pp. 371 and 663), repetition of material (pp. 55–57 and 115–116) or narrative-like history provided by some researchers when describing their own work.

Smaller errors, misinterpretations or confusing remarks can be found, as in equation 1 (p. 173), where the surface tension γ is misprinted as g ; equation 9 (p. 200), which is dimensionally incorrect; the oversimplification of the nonlinear viscoelastic behaviour of materials (notably the modulus definition on p. 205); the one-sided statement about methods to characterize protein adsorption (p. 212), which excludes ATR-FTIR, circular dichroism and radioactive labelling techniques; and the assignment of finite molecular weights to insoluble crosslinked gels (p. 681).

Nevertheless, the book achieves its main goal of educating and directing the novice and advanced researcher in the field. □

Nicholas A. Peppas is at the School of Chemical Engineering, Purdue University, West Lafayette, Indiana 47907-1283, USA.

The living and the dead

The Evolutionary Biology of Plants

by Karl J. Niklas

University of Chicago Press: 1997. Pp. 449.
\$65, £51.95 (hbk); \$19.95, £15.95 (pbk).

Bill Chaloner

There is a cultural divide in the spectrum of scientists who study plant evolution. At one end are those interested in the record in the rocks of what happened to plants over geological time. They base their work in the field, searching for the remains of past plant life in boreholes, mines and quarries. And having investigated these remains in the laboratory, they then try to reconstruct the biology of the plants as they were in life and interpret the way in which these plants have changed down the ages.

At the other end of the spectrum are those who seek to understand not so much what happened as how it happened. Some tackle this at the molecular and cellular level, looking at the genetic basis of inheritance and the aberrations that produce evolutionary change. Others with similar objectives study living plants and plant populations to understand the evolutionary process in action.

Differences of perspective and academic background have kept the palaeobotanist and the evolutionary geneticist apart. Karl J. Niklas now makes an inspired attempt to bridge this gap. He takes us through the current perceptions of the workings of plant genetics and the diversity of adaptation produced by natural selection and then looks back at the fossil record. Finally, he has a good shot at trying to bring the two fields together. One of his stated aims in writing the book is to "give the living and the dead equal attention", and this he largely achieves.

Much of Niklas's scholarly review of the origin of life, the evidence from cell biology of the relationship between prokaryotes and eukaryotes, and the processes of inheritance, mutation and evolution have been well presented in sources elsewhere. These are intended to educate the palaeobotanists. The important contribution of this book is to present to the other side of the cultural divide the evidence of plant evolution that has come from studying fossil plants during the past two or three decades.

But this is much more than a broad-based overview of palaeobotany. Niklas brings to the story a diverting assembly of new dimensions, many of which he has contributed to palaeobotany by his own research. He explores with simple computer models how plant architecture, leaf shape and arrangement have moulded plant evolution, as plants struck compromises between these structures and various environmental constraints. These "walks on a

fitness landscape" (a phrase borrowed from Sewall Wright) bring us back, but with a new insight, into problems of size and form in plants that were being explored by F. O. Bower 70 years ago.

Niklas's own research has included such diverse items as explaining how the properties of plant tissues can help us to understand the mechanical designs of extinct plants with their very different patterns of development from any now living; how the shape of early wind-pollinated seeds contributed to their aerodynamic efficiency; and how the pattern of plant evolution works on different rules from the innovations and mass extinction of the animal kingdom.

He also offers interesting new perspectives on many old themes. Insect pollination, which lies at the heart of flowering plant biology, works only "because animals can recognise flowers belonging to the same

species". In other words, insect pollination paid off because plants could teach insects to recognize different shapes and colours and so to adopt brand fidelity (a vital commercial principle — we pick the same package off the supermarket shelf each time). But it worked only because the insect brain could acquire a sound taxonomic training at the hands of the flowers.

In the final chapter, Niklas comes into his own in integrating the dead with the living, as he explores the molecular clock of plant evolution, the pattern of evolutionary tempo, and the mysteries of mass extinctions. Plant biologists have in one sense been waiting many years for this book, and yet perhaps its real importance lies in the fact they did not realize it was coming. □

Bill Chaloner is in the Department of Geology, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK.

Other books on plants

Diversity and Classification of Flowering Plants

by Armen Takhtajan

Columbia University Press, \$109, £76

A new revision of the author's 1967 classification system based on "descriptive keys" for individual families.

Mycorrhizal Symbiosis

by S. E. Smith and D. J. Read

Academic, \$74.95, £65

Second edition of a work that first appeared in 1983, providing a comprehensive review of the symbiotic fungi that colonize plant roots.

New Flora of the British Isles

by Clive Stace

Cambridge University Press, £28.95, \$85 (pbk)

The second edition of this standard work now covers more than 4,500 taxa.

Plant Ecology

edited by Michael J. Crawley

Blackwell Science, £29.95 (pbk)

Second edition of this benchmark text, first published in 1986.

Plant Functional Types: Their Relevance to Ecosystem Properties and Global Change

edited by T. M. Smith, H. H. Shugart and F. I. Woodward

Cambridge University Press, £80, \$120 (hbk);

£29.95, \$44.95 (pbk)

How to predict the effects of changing climate and carbon dioxide on plants at the global scale.

Seed Production: Principles and Practices

by Miller B. McDonald and Lawrence Copeland

Chapman and Hall, £49

Reference text for students.

Plant Genetic Conservation: The In Situ Approach

edited by N. Maxted, B. V. Ford-Lloyd

and J. G. Hawkes

Chapman and Hall, £27.50 (pbk)

Includes methodologies and detailed discussion of on-farm and genetic-reserve conservation.

Reaching for the Sun: How Plants Work

by John King

Cambridge University Press, £30, \$54.95 (hbk);

£9.95, \$16.95 (pbk)

An accessible account for lay people.

Practical Applications of Plant Molecular Biology

by R. J. Henry

Chapman and Hall, £24.99 (pbk)

Introductory textbook and reference source.

World Weeds: Natural Histories and Distribution

by LeRoy Holm, Jerry Doll, Eric Holm, Juan

Pancho and James Herberger

Wiley, £150

Comprehensive coverage of more than 100 species, with a bibliography of 3,000 references.

The Identification of Flowering Plant Families

by James Cullen

Cambridge University Press, £21.95, \$59.95 (hbk);

£12.95, \$35 (pbk)

Fourth edition of a short, user-friendly guide aimed at botany and horticulture students.

The Nature of Disease in Plants

by Robert P. Scheffer

Cambridge University Press, £45, \$64.95

A natural history of some of the most damaging diseases worldwide.

Crystal structures of fragment D from human fibrinogen and its crosslinked counterpart from fibrin

Glen Spraggon*, Stephen J. Everse* & Russell F. Doolittle

Center for Molecular Genetics, University of California, San Diego, La Jolla, California 92093-0634, USA

* These authors contributed equally to this work.

In blood coagulation, units of the protein fibrinogen pack together to form a fibrin clot, but a crystal structure for fibrinogen is needed to understand how this is achieved. The structure of a core fragment (fragment D) from human fibrinogen has now been determined to 2.9 Å resolution. The 86K three-chained structure consists of a coiled-coil region and two homologous globular entities oriented at approximately 130 degrees to each other. Additionally, the covalently bound dimer of fragment D, known as 'double-D', was isolated from human fibrin, crystallized in the presence of a Gly-Pro-Arg-Pro-amide peptide ligand, which simulates the donor polymerization site, and its structure solved by molecular replacement with the model of fragment D.

Fibrinogen units pack together to form a fibrin clot¹, but without a crystal structure it is difficult to understand how this is achieved. Fibrinogen is readily purified from blood plasma, but it is fragile, and efforts to crystallize the native molecule have so far been unsuccessful. But electron microscopy long ago revealed that fibrinogen has a trinodular structure², the distal globular entities of which were thought to be joined to a central domain by 'coiled coils'³. Biochemical characterization showed the molecule to be a covalent dimer composed of six polypeptide chains ($\alpha_2\beta_2\gamma_2$) linked by an extensive network of disulphide bonds, and enormous effort had to be expended to understand how the six chains were accommodated by the triglobular structure. It turns out that each of the two terminal globular domains contains the three non-identical chains ($\alpha\beta\gamma$), and that these are tethered together in the smaller central domain, which contains all six amino termini. The three-stranded connectors can be broken by proteases such as plasmin, generating two principal core fragments: a small fragment corresponding to the central domain (fragment E), and a set of larger fragments corresponding to terminal portions (fragments D). The two D fragments together account for half the mass of the native molecule. In contrast, fragment E, which is mostly composed of the two three-stranded connectors joined near their N termini, amounts to only 15% of the starting molecule. The remaining third of the protein is accounted for by the free-standing C-terminal two-thirds of the α -chains, which are readily removed by most proteases.

It has been presumed that both fragments D and E contain substantial portions of the severed 'coiled coil', the boundaries of which have been thought to be 'disulphide rings'⁴. Thus the three non-identical chains in each half of a fibrinogen molecule are interconnected by two sets of 'disulphide rings', which are distinguished by homologous braces of cysteines separated in each case by three residues.

Numerous biochemical and biophysical studies have been conducted on fragments derived from both fibrinogen and fibrin in an effort to understand the formation of fibrin clots and fibrinolysis. Briefly, clotting occurs when thrombin removes fibrinopeptides from the centrally located α -chain N termini, thereby exposing new N-terminal sequences known as 'knobs' (in the fragment E region). These spontaneously interact with 'holes' on the distal domains of other fibrinogen molecules (in the fragment D regions). Because the molecule is a dimer, oligomerization can proceed in either direc-

tion, the result being a two-molecule thick protofibril with a half-molecule staggered overlap. Peptides fashioned on the 'knob' sequences bind to fibrinogen, as well as to fragment D, and can effectively prevent polymerization⁵. It is because fragment D contains the 'hole' half of the knob-hole interaction that it is physiologically active as an inhibitor of fibrin formation.

Concomitantly, the polymerizing fibrin oligomers are reinforced by the thrombin-activated transglutaminase known as factor XIII, which introduces γ -glutamyl- ϵ -amino-lysine isopeptide crosslinks between the carboxyl-segments of abutting γ -chains⁶. The protofibrils subsequently associate laterally to form first fibrils and then mature fibres, additional crosslinks slowly being added between the carboxyl regions of a small fraction of α -chains. When fibrin formed under crosslinking conditions is proteolysed, the core fragments released are E and crosslinked D ('double-D' or 'D-dimer')⁷, as well as the unstable complex known as 'D₂E'⁸.

Although a detailed atomic structure of fibrinogen has not yet been determined, combined electron microscopy and X-ray diffraction efforts have yielded an 18-Å structure for a protease-modified molecule⁹, and recently a high-resolution X-ray structure has been reported for one of the two domains of M_r 30K that constitute the terminal globules¹⁰. We now report an X-ray structure at 2.9 Å for

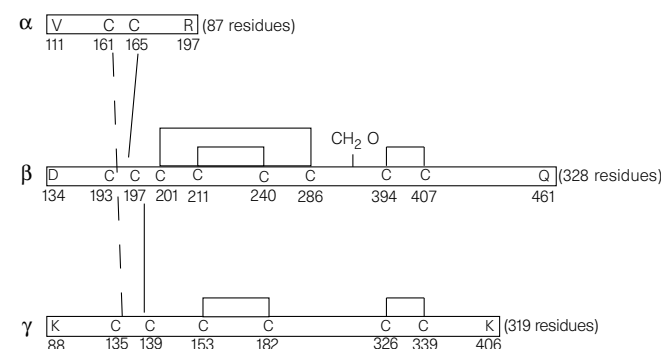


Figure 1 Schematic representation of the three polypeptide chains that compose human fibrinogen fragment D. The 16 cysteine residues are involved in three interchain and five intrachain disulphide bonds¹⁴. Residue numbering corresponds to native human fibrinogen, in which the α -chain runs from Ala¹ to Val⁶¹⁰, the β -chain from PCA¹ to Gln⁴⁶¹, and the γ -chain from Tyr¹ to Val⁴¹¹.

the 86K entity known as fragment D. It is composed of three disulphide-linked polypeptide chains ($\alpha'\beta'\gamma'$) and contains 734 amino-acid residues, the α' , β' and γ' chains having 87, 328 and 319 residues, respectively (Fig. 1).

We have also isolated the factor XIII-crosslinked counterpart, double-D, from a carefully controlled proteolytic digest of human fibrin, and solved its structure by molecular replacement with the model of fragment D. Moreover, the double-D was co-crystallized with the ligand Gly-Pro-Arg-Pro-amide, which mimics the 'knobs' contributed by the fragment E portion of the knob-hole interaction⁵. As a result, the structure allows us to determine the three main events that occur during the initial stages of fibrin formation: the Gly-Pro-Arg knob-hole interaction; the D-D 'self-association'; and the whereabouts of the γ - γ crosslinks. Finally, given the locations of the surrogate 'knobs' and a general knowledge of the extent of the coiled coils and the dimensions of fragment E, we were able to construct a model of the fundamental fibrin monomer and the oligomeric protofibril.

Structure determination

The structure of fragment D was solved by the use of multiple isomorphous replacement and anomalous scattering (MIRAS) techniques, using gold, mercury and uranium derivatives. The initial MIRAS phases were improved by solvent flattening, after which tubes of electron density corresponding to the expected coiled coils could be discerned in a 6-Å map; a clear delineation of the entire monomer envelope by inspection was not possible, however. The matter was resolved by an envelope-skeletonization

procedure (G.S., in preparation), which provided a preliminary mask that was subsequently used as a search model with the self-rotation angles to determine accurate non-crystallographic symmetry (n.c.s.) matrices for the structure. The initial mask was improved by correlation mapping, after which averaging was performed on the maps themselves, followed by limited phase extension to 3.5 Å. This strategy resulted in a considerably improved map that allowed almost an entire chain trace of the molecule to be made. After the initial trace was made, the quality of maps was improved by cross-averaging the native fragment-D map with that from a non-isomorphous crystal of fragment D and maps from double-D (below). During the model-building phase, we took advantage of the homologous nature of the β and γ globular domains, the better density in one or the other of these serving as a guide.

The structure was refined using a combination of positional, simulated annealing, and thermal parameter refinement, beginning from a crystallographic *R*-value of 52.1% at 2.9-Å resolution. This was followed by alternating cycles of manual refitting to SigmaA-weighted and cross-averaged maps. The model refinement was monitored closely by determining the free *R*-factor, and all weights were set on this basis. The model was refined with n.c.s. constraints throughout. In the latter stages of refinement, an anisotropy and bulk solvent correction were included. In the end, a model with good statistics at 2.9 Å emerged (Table 1); the conventional *R*-factor was 26.3% and the free *R*-factor 36.3%, with good stereochemistry as shown by Ramachandran plots. Only two residues involved in disulphide bonds fell into disallowed regions (Cys β 407, Cys γ 339). In total, 707 of the chemically expected 734 residues were positioned; the 27 absences were all situated at the termini: β 134–147, α 196–197, β 461 and γ 397–406. In the final stages of building and refinement, the electron density improved to the point where the single calcium atom¹¹ could be added to the γ -chain globular domain, as was a nominal sugar residue in the case of the β -chain¹².

The structure of the factor XIII-crosslinked double-D was determined by molecular replacement using an unrefined chain trace of native D as a model; all data in the resolution range 8–4 Å were used. The rotation search generated two peaks approximately 180 deg apart, both of which were subjected to two-dimensional translation searching. The correlation coefficients for the two solutions were 22.8% (4.8 σ above the mean) and 24.9% (7.5 σ above the mean), respectively; the corresponding *R*-factors were 52.9% and 52.4%. When one solution was fixed and the other searched for in the *y*-direction for the four origins of the *P*₂₁ asymmetric unit, the correlation coefficient rose to 38.7% (the next highest solution was 26.3%) and the *R*-factor dropped to 48.4%. The electron density of the Gly-Pro-Arg-Pro-amide ligands was clearly evident in a $2|F_o| - |F_c|$ map.

The double-D structure was refined beginning from a crystallographic *R*-factor of 50.9% on all data between 8.0 and 2.9 Å resolution. This was followed by alternating cycles of manual refitting to $2|F_o| - |F_c|$ maps and refinement of atomic positions

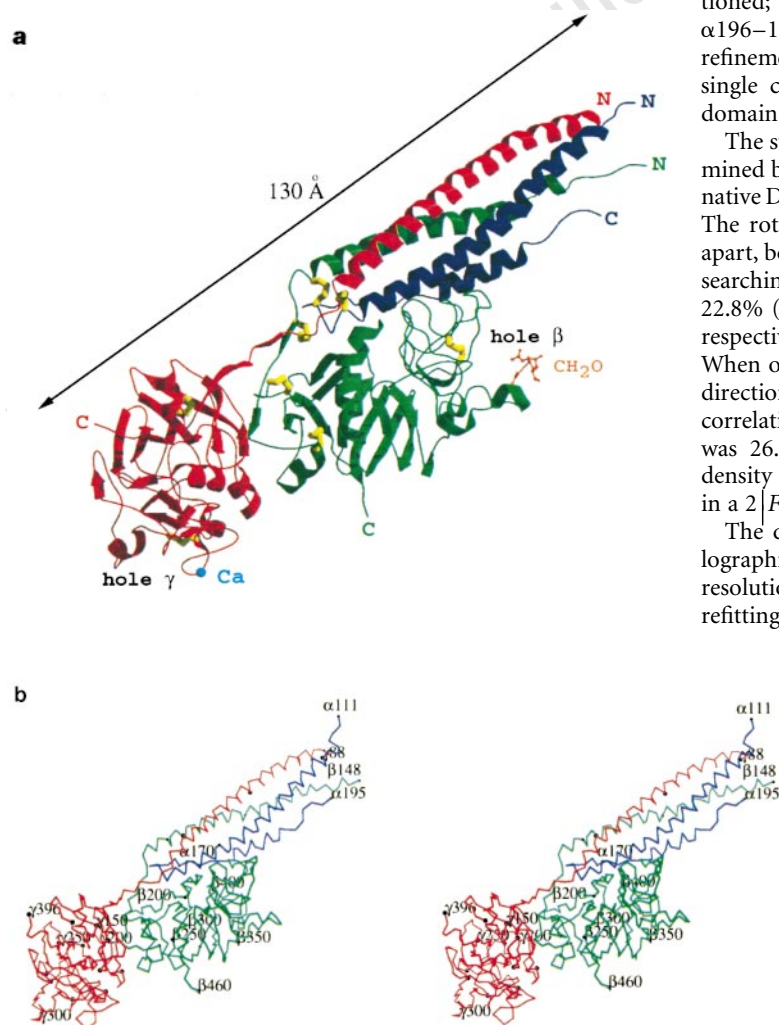


Figure 2 **a**, Ribbon representation of fragment D, showing the region of coiled coils and the globular β and γ domains. Colour scheme: blue, α -chain; green, β -chain; red, γ -chain; yellow, disulphide bonds. N and C denote amino and carboxy termini, respectively; Ca, calcium-binding site; CH₂O, carbohydrate cluster at Asn β 364; 'hole' denotes potential binding cavity. **b**, Stereo view of α -carbon trace of fragment D. Markers on chains are shown at 50-residue intervals.

and *B*-factors. The final conventional *R*-factor was 24.1% and the free *R*-factor was 31.8% with good stereochemistry (Table 1). Refinement was monitored closely by the free *R*-factor, and all weights were determined on this basis. Initially, the model was refined with the use of n.c.s. constraints. In later stages of refinement, a restrained regime was adopted, warranted by a significant drop in the free *R*-factor upon application of weights of 100 kcal mol⁻¹ imposed on all atoms related by n.c.s. In the very last stages, an anisotropic *B*-factor and bulk solvent correction were applied, and all data from 30 to 2.9 Å were included.

General description of fragment D structure

The overall architecture of the 130-Å long fragment is plough-shaped, beginning with the three-stranded coiled-coil region and a double-back fourth helical segment of α -chain, and continuing on to two prominent globular regions corresponding to the carboxyl halves of the β - and γ -chains, respectively (Fig. 2). The two domains are orientated relative to each other at about 130 deg. The β domain is folded back close to the coiled coil; the γ domain extends onwards to a distal position.

The remnant coiled coil in fragment D consists of residues Val α 111 to Ser α 160 (50 residues), Asp β 134 to Tyr β 192 (59 residues), and Lys γ 88 to Gln γ 134 (47 residues). The N-terminal 14 residues of the β -chain, which are mostly unmatched in the other chains (Fig. 1), are disordered and do not appear in the model. The

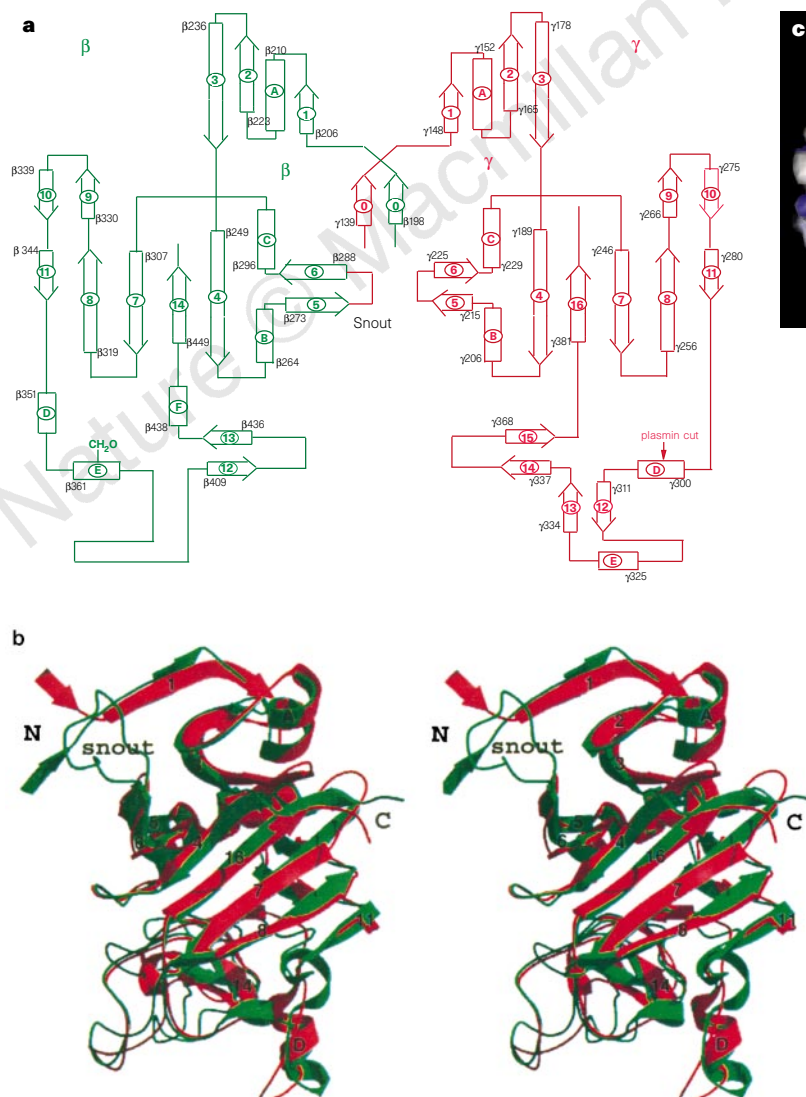
packing of the three chains in the coiled coil is canonical, with most of the non-polar side chains directed inward. The clockwise arrangement of the three chains in fragment D, as seen from the N termini, is α - γ - β .

The fourth helix in the bundle corresponds to residues Ser α 166 to Pro α 195. It nestles in the groove between the α - and β -chains on the face of the coiled coil opposite to the γ -chain with its non-polar side chains directed inward. The situation is reminiscent of the influenza haemagglutinin HA₂ structure, which also has a hairpin helix associated with a coiled coil¹³.

The disulphide connections linking the three chains are different from what was expected. Previous chemical determinations¹⁴ had established that Cys β 193 is linked to Cys α 165, but the other two disulphides could not be assigned at that time. It was predicted on symmetry grounds, however, that they would connect Cys α 161 to Cys γ 139 and Cys γ 135 to Cys β 197. Instead, the X-ray structure reveals that, whereas Cys β 193 is indeed connected to Cys α 165, the other two connections are between Cys α 161 and Cys γ 135, and Cys γ 139 and Cys β 197. The asymmetric arrangement is in keeping with the unexpected reversal of direction for the α chain.

Homologous globular domains

The β - and γ -chains form a short two-stranded parallel β -sheet (strands denoted '0' in Fig. 3a) as they proceed away from the disulphide junction before adopting separate courses that orientate



them at about 130 deg to each other. Beyond that, there is only one interaction between the two domains, an antiparallel β -sheet composed of residues Asn β 202–Val β 205 and Gly γ 216–Sery γ 219. Each of the homologous domains consists of two small subdomains, at the N- and C-terminal ends, respectively, and a central domain with a 5-stranded antiparallel β -sheet (Fig. 3a). A large extended helix is interrupted by a snout-like region. In the β -chain, the snout-like region protrudes further than in the γ -chain, the result of an eight-residue insertion that includes Cys β 286, which is bonded to Cys β 201 near the disulphide junction.

In both chains, the C-terminal subdomain is composed of three extended loops that form the cavity expected for a ligand-binding site. As anticipated, the calcium-binding site in the γ domain is contiguous to the binding cavity (Fig. 2), as is the protease-sensitive bond at Lys γ 302. One unexpected feature of the structure is that residues Lys γ 381–Asp γ 390 in one case, and Lys β 453–Phe β 458, in the other, fold back from the C-terminal subdomain and form a central β strand in the main domain.

For the most part, our structure is in good agreement with a recently reported structure of the γ -chain globular domain¹⁰. In that

Table 1 Statistics for data collection, phase determination and refinement

	Fragment D						Double D	
	Space group: $P2_1$						Space group: $P2_1$	
	Unit cell: $a = 107.72, b = 48.08, c = 167.56, \beta = 105.70^\circ$						Unit cell: $a = 93.82, b = 95.5, c = 113.76, \beta = 96.08^\circ$	
			Asymmetric unit: 2 molecules				Asymmetric unit: 1 molecule	
	Native	KAuCl ₄	KAuCl ₄	HgI ₄	HgI ₄	UO ₂ (NO ₃) ₂		
			(anomalous)	(anomalous)	(anomalous)			
Beam	UCSD	Daresbury	Daresbury	Daresbury	SSRL	Daresbury	SSRL	BNL
Detector	MW	Mar	Mar	Mar	Mar	Mar	Mar	Mar
λ (Å)	1.54	0.88	0.88	0.88	1.08	0.88	1.08	1.02
$d_{\min}$ (Å)		2.9	2.9	2.9	3.3	2.9	3.3	2.9
Number of crystals		9	1	1	1	1	1	5
Observations (N)		318,238	141,315	141,315	140,160	90,389	110,282	244,386
Unique reflections (N)		35,552	27,080	27,074	24,106	19,596	23,548	40,637
Mean redundancy		8.95	5.2	5.2	5.8	4.6	4.7	6.0
Data coverage (%)		87.1	71	71	92	60	81	88.8
R_{sym} (I) (%)		14.5	9.5	9.5	18.2	12.5	13.2	19.2
Mean isomorphous difference (%)†			18.0			26.1		14.6
MIR analysis								
$d_{\min}$ (Å)			4.0	4.0	5.0	5.0	4.5	
Heavy-atom sites/asymmetric unit			4	4	1	1	2	
Phasing power‡			1.72	1.71	0.83	1.27	1.16	
$R_{\text{cullis}}§$			0.607		0.693		0.658	
Mean overall figure of merit		0.496						
Refinement								
		Native D model: 707 residues (5,712 atoms)				DD Model: 1,414 residues (11,391 atoms)		
d -spacings (Å)		30.0–2.9				30.0–2.9		
R -value¶ (%)		26.3				24.1		
Free R -value		36.3				31.8		
r.m.s.d. bonds (Å)		0.011				0.010		
r.m.s.d. angles (°)		1.63				1.54		
r.m.s.d. B values (Å ²)		4.7				4.3		
Average B value		25.00				30.81		

* $R_{\text{sym}} = \sum |I - \langle I \rangle| / \sum I$.
† Mean isomorphous difference = $\sum |F_{(P)} - F_{(PH)}| / \sum F_{(P)}$, where $F_{(P)}$ is the protein structure factor amplitude and $F_{(PH)}$ is the heavy-atom derivative structure factor amplitude.
‡ Phasing power = r.m.s. $\sum |F_H| / E$, where F_H is the calculated heavy-atom structure factor amplitude and E is the lack of closure error.
§ $R_{\text{cullis}} = \sum ||F_{(PH)}| - |F_{(P)}|| / \sum |F_{(PH)}|$ for centric reflections only.
|| Mean figure of merit = $\sum |F_{(obs)}| - |F_{(calc)}| / \sum |F_{(obs)}|$.
¶ Crystallographic R -value = $\sum ||F_{(obs)}| - |F_{(calc)}|| / \sum |F_{(obs)}|$ with 95% of the native data for refinement. Free R -value: R -value based on 5% of the native data withheld from refinement.
MW, multiwire detector; Mar, Mar imaging plate; SSRL, Stanford Synchrotron Radiation Laboratory.

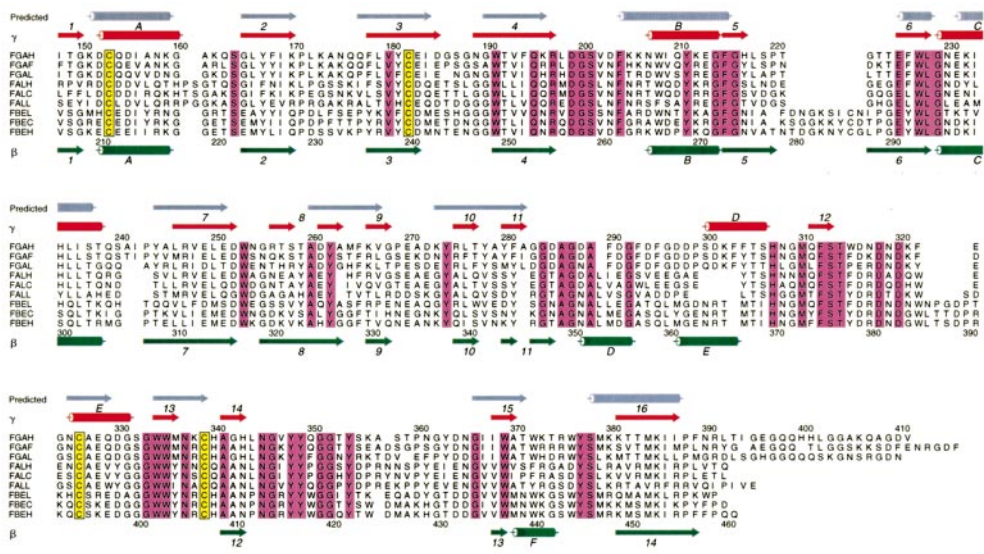


Figure 4 Multiple alignment of carboxyl domains of fibrinogen β (FBE)- and γ (FGA)-chains from various species (H, human; F, frog; C, chicken; L, lamprey). Also included are related domains from minor forms of (FAL) chains^{49,50}. The observed secondary structures of the human γ - and β -chains, with their human numbering, are shown above and below the alignment, respectively; α -helices are denoted with capital letters and β strands with numbers. A long-standing secondary-structure prediction¹⁵ is shown in grey at the top.

study, some ambiguity was reported for the C-terminal segment, different crystal forms giving rise to different conformations. Similarly, in our native fragment-D structure we were not able to model in the last 10 residues at the C-terminal end of the γ -chain. It was possible, however, to trace the shorter β -chain almost all the way to the C-terminal Gln β 461 in the native fragment D. For the most part, the β - and γ -chain folds are very similar (Fig. 3b), the root-mean-square (r.m.s.) distance being 1.42 Å for 220 aligned residues.

There are a few important differences, however, including two short helices in the β -chain and a short helix and some additional β structure in the γ -chain (Figs 3a and 4). Also, the β domain has a five-residue insertion that, as forecast¹⁵, disrupts the calcium-binding site. In addition, the β -chain has a carbohydrate cluster at Asn β 364 (Fig. 1), which is reported to be of the biantennary type¹². The non-protein density in this region is disordered, and only a nominal sugar was incorporated in the model.

As noted above, both the β and γ domains have large cavities that are well disposed for binding peptides (Fig. 3c). Equilibrium dialysis studies⁵ have shown that both sites should bind Gly-Pro-Arg-peptides, but only the β site should bind Gly-His-Arg-peptides, which correspond to the thrombin-exposed N termini of β -chains. The binding cavity in the γ -chain has much more electronegative quality than does the β -chain equivalent (Fig. 3c). It may also be important that the single sugar residue incorporated in the model is positioned at the very mouth of the β -chain cavity. In the γ -chain, the structures in good accord with the biochemistry, Tyr γ 363 being inside the γ -chain binding cavity, as expected on the basis of affinity-labelling studies¹⁶.

A search of reported three-dimensional structures did not show any significant resemblances to known proteins, a result in accord with the recent report on the γ -chain carboxyl domain¹⁰. At the sequence level, however, the β and γ globular domains are known to be homologous with numerous other C-terminal domains found in animal extracellular proteins, including tenascins, restrictins, ficolin and calcium-dependent lectins found in both vertebrates and invertebrates. Previously, an effort had been made to predict the secondary structure of these domains on the basis of a diverse multiple alignment and various biochemical considerations¹⁵. If a three-state criterion is used (α -helix, β -sheet or other), then 81% of

the residues were assigned correctly (Fig. 4). As anticipated, the accuracy of the prediction was greater in the N-terminal half of the domain, the looser carboxyl-segment being more ambiguous. Another prediction with much the same results has recently appeared¹⁷.

Structure of double-D

In general, the structure of the double-D closely resembles that of two molecules of fragment D butted end-to-end, conformational changes as such being modest. There are some slight changes in the positions of the helices in the coiled coil near their N termini, and the ligand-binding pocket in the γ -chain does undergo some readjustment. In this regard, the binding of the Gly-Pro-Arg 'knobs' involves both electrostatic interactions and hydrogen bonding, the α -amino group being juxtaposed between residues Asp γ 364 and Asp γ 330, and the guanidino group of the arginine lying between the carboxyl group of Asp γ 364 and Gln γ 329 (Fig. 5a, b). Variant human fibrinogens with defective polymerization have been reported for all three of these sites^{18,19}. Fibrinogen Matsumoto I has Asp γ 364 mutated to His, fibrinogens Kyoto III and Milano have Asp γ 330 mutated to Tyr and Val, respectively, and fibrinogen Nagoya has Gln γ 329 changed to Arg. Some other reported variants are involved less directly by disrupting the network of hydrogen bonds involving important residues. Additionally, the loop containing residue Tyr γ 363, which is known to be within the binding cavity on the basis of affinity-labelling experiments¹⁶, has moved relative to its position in native D. The r.m.s.d. for the C α atoms of the segment Lys γ 356–Asn γ 365 is 1.54 Å.

The two γ -chain binding cavities are on the same side of the fibrin dimer, which has a 'sidedness' or dorsal–ventral aspect (Fig. 6a, b). The fragment E region of fibrinogen, which includes the two Gly-Pro-Arg 'knobs' in the native molecule and is intrinsically dimeric, must straddle the junction on the side diametrically opposite the crosslinks. Because the two α -chains contributing the Gly-Pro-Arg 'knobs' (residues Gly α 17–Arg α 19) are joined by a disulphide bond only nine residues away (Cys α 28) that necessarily falls on the two-fold axis of native fibrinogen, it was possible to position a model mask of the fragment E as it would be in the 'D₂E' complex and the beginning protofibril (Fig. 6c, d). In this regard, a crude model of fragment E was constructed by first applying a rotation of 180 deg

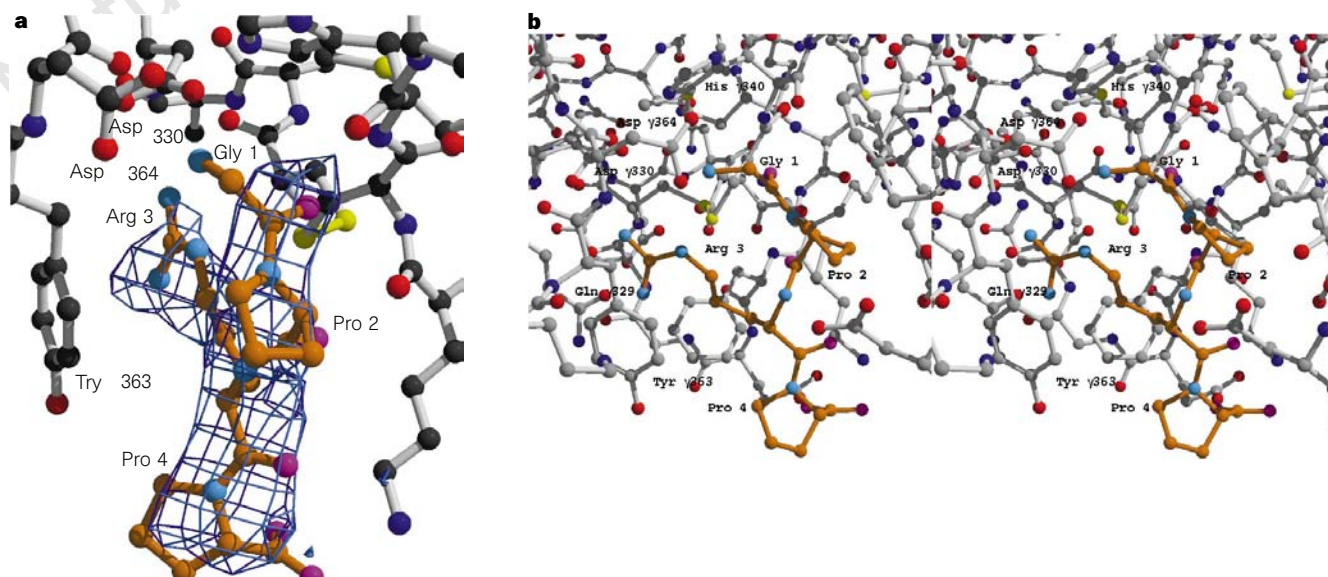


Figure 5 a, Electron-density omit map showing peptide ligand Gly-Pro-Arg-Pro-amide bound to γ -chain binding site in double-D as calculated with $|F_o| - |F_c|$ coefficients and phases from the refined model contoured at 2.2σ . The ligand

model was not included in the F_o calculation. **b**, Stereoscopic view of ball-and-stick model showing ligand position in binding site. Residue numbering corresponds to native human fibrinogen.

along the long axis of double-D, followed by translations of the half-stagger distance (225 Å) along this same axis and the width of double-D in a perpendicular direction. The coiled coil of this reoriented double-D was extended backwards by the correct number of residues to reach the disulphide ring in fragment E

(for example, Cys α 45–Cys α 49). It was then a simple matter to draw the short connections to the disulphide dyad (Cys α 28–Cys α 28) and on to the experimentally determined ‘knobs’ in the untranslated double-D. The other half of fragment E was constructed by applying the n.c.s. operators of the double-D structure. The symmetrical model for fragment E is in the form of a double dog-leg, the result of being anchored on the two ligands on the one hand, and the extensions to meet the coiled coils in D on the other (Fig. 6c). A model of the protofibril can be constructed by repeated application of the operations and their inverse (Fig. 6d).

End-to-end association of fibrin monomers.

The end-to-end association of joined fragments D occurs in a staggered fashion such that a significant but not maximal amount of surface area is occluded between the extremities (Fig. 6a). Accessibility measurements indicate that approximately 750 Å² of surface is lost as a result of the association, a modest contribution to the free energy of polymerization²⁰. There is a distinct crevice between the associated ends that probably contains some solvent. The offset nature of the association relative to the major molecular axis requires an asymmetric set of interactions between the adjacent fibrin units. Of particular interest is residue Arg γ 275, one of the most frequently reported mutations in polymerization-defective human fibrinogens¹⁹. In native fragment D, this residue has its guanidino group within hydrogen-bonding distance of the carboxyl

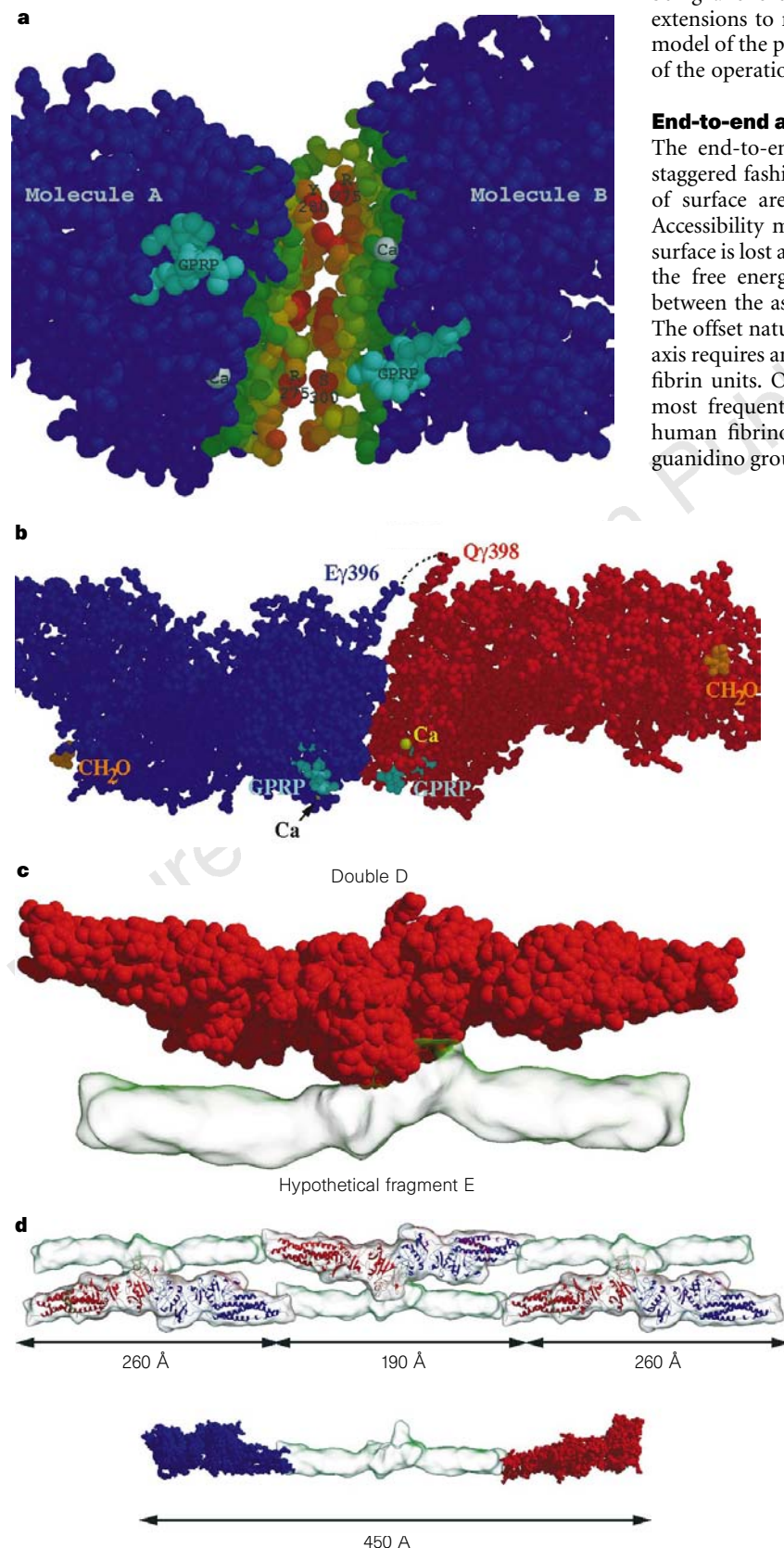


Figure 6 Four views across the D-D interface as seen from increasing distance. **a**, Close-up showing structural details of end-to-end abutment of fragments D in double-D: R, Arg γ 275; Y, Tyr γ 280; S, Ser γ 300; Ca, calcium atoms; GPRP, Gly-Pro-Arg-Pro-amide ligands. The residues at the interface are coloured according to their nearest distances to residues on the other side of the crevice, red being shortest and greater distances shading toward green. **b**, More distant view showing Glu γ 396 of molecule A and Gln γ 398 of molecule B connected by a broken line denoting crosslink; CH₂O, carbohydrate. **c**, Even more distant view of fibrin double-D with hypothetical fragment E domain rooted on the two Gly-Pro-Arg ligands and straddling the interface of abutting D domains. **d**, Very distant view showing three molecules of double-D connected by hypothetical structures corresponding to full-sized fragments E and connecting segments to show required packing in protofibrils. A fibrin monomer made from two halves of the double-D structure and the hypothetical fragment E is shown at the bottom.

group of Asp γ 272. The two corresponding arginine residues in double-D are not significantly moved in that regard, but two different intermolecular interactions are now evident. In molecule A, the guanidino group is within hydrogen-bonding distance of the hydroxyl group of Ser γ 300 on molecule B. The equivalent group on molecule B is opposite Tyr γ 280 on molecule A.

Although calcium has been reported to improve fibrin formation in some mutant fibrinogens¹⁹, there is no evidence that calcium is involved in the end-to-end association, as revealed by the structure, and the effect must either be the result of maintaining the general structural integrity at the nearby calcium-binding site (Fig. 6a, b), or the calcium is contributing to a later stage in polymerization.

It has long been recognized that fibrin formation involves an end-to-end interaction independent of either the Gly-Pro-Arg-dependent knob-hole interaction or the γ - γ crosslinks, and it has been predicted²¹, on the basis of the variant human fibrinogen denoted Tokyo II, that Arg γ 275 must be involved in that end-to-end interaction. The X-ray structure verifies that prediction exactly.

Crosslinks

The reciprocally linked γ -chain carboxyl segments are not visible in the electron-density maps and must have some mobility despite being connected from one unit to another. The logic of their whereabouts derives from the observations that the two fragments D are chemically linked to each other, as is readily shown on SDS gels, and that the X-ray structure shows that the two fragments D are situated end-to-end. The ϵ -amino γ -glutamyl crosslinks, which are known⁶ to involve Gln γ 398 and Lys γ 406, join two molecules that are end-to-end. Because of the offset nature of the end-to-end interaction, the two different γ -chain carboxyl segments must have different conformations (Fig. 6b). This may explain why different investigators have reported different conformations for this region^{10,22}. In our maps, the electron density was discernible up to Glu γ 396 in molecule A, and up to Gln γ 398, the crosslink acceptor residue, in molecule B (Fig. 6b). The α - α distance between them is 17 Å, perhaps slightly too far for a simple helical connector of the 10 residues (15 Å) needed to reach the molecule A donor at Lys γ 406, but definitely too short for a simple antiparallel β strand (34 Å). In any case, the connections are on the opposite side of the joined units from where the knob-hole interactions occur.

There has long been debate over whether the γ - γ crosslinks connect neighbouring molecules in an end-to-end manner (longitudinal)²³ or side by side (transverse)²⁴. The crystal structure now demonstrates that they are connected end to end. Nonetheless, the crosslinked segments are so exposed and readily accessible to factor XIII that it should be possible for additional crosslinks to be formed with other protofibrils, which may explain reports of crosslinked trimers and tetramers²⁵. The peripheral location of the crosslinks also shows how the extended γ -chain known as γ' , reported to be a binding site for factor XIII²⁶, would remain external after the association of fibrin monomers and would itself be cross-linked.

Other physiological considerations

As well as providing details about how the individual units are packed in the protofibril, these structures allow us to investigate subsequent interactions in fibrin formation. They also allow us to consider other important differences between fibrinogen and fibrin, particularly the physiological matter of how various cells and macromolecules distinguish between these two entities.

The reversal of the α -chain at the fragment-D disulphide ring came as a surprise, but in retrospect it may make good biological sense for several reasons. First, the known plasmid cut-points in the region Lys α 206-Arg α 239 are near the cuts that sever the coiled coil. Second, the direction of the chain in this region is consistent with reports that the C-terminal domains of α -chains are located in the central parts of the molecule²⁷. Third, the model in which the

intramolecular associations of α -chain C-terminal domains are broken during the polymerization process and reassociated in an intermolecular mode²⁸ might explain how certain features of the coiled coil could be exposed in fibrin but not in fibrinogen.

As an example, the sequence Lys α 151-Arg α 159, which is part of the coiled coil immediately adjacent to the disulphide ring, has been reported to activate tissue-type plasminogen activator in fibrin but not in fibrinogen²⁹. In fragment D, this region of the α -chain is closed off on one side by the β globular domain, on two others by the β - and γ -helices, and on its 'frontside' by the fourth helix. Measures of solvent accessibility showed the segment to be totally buried. Given the proposal that the α -chain carboxyl region moves away from the parent molecule during polymerization²⁸, it might seem reasonable that the fourth helix could be pulled away from the coiled coil and expose the masked residues, but we observed no significant difference in this region in the double-D fragment isolated from fibrin. One possibility is that the fourth helix is pulled away, but then snaps back into the groove after plasmin clipping, in which case the double-D would not be expected to be an activator of t-PA, but it has been reported to be active³⁰. For the moment, the important difference between fibrin and fibrinogen with regard to biologically active groups remains unresolved.

A recent report shows that the Gly-Pro-Arg-Pro-amide ligand binds to the 30K γ -chain domain³¹. □

Methods

Protein preparation and crystallization. Fragment D was prepared from trypsin-digested human fibrinogen³² made from bloodbank plasma. N-terminal sequencing revealed an almost homogeneous protein with Val α 111, Asp β 134 and Met γ 89, with a small amount of Lys γ 88 present. We thank T. Takagi of Tohoku University for determining the sequences. Sialic acids were removed by digestion with neuraminidase; some preparations were purified further by DEAE-chromatography. Crystals were grown at room temperature from sitting drops with wells containing 16–19% PEG 3350, 50 mM Tris, pH 7.5–8.5, 2 mM Na $_3$ N and 50–133 mM CaCl $_2$. Macroseeding was used to obtain large crystals for diffraction studies³³.

The double-D fragment was prepared directly from lysed crosslinked fibrin generated by adding bovine thrombin (0.1 U ml $^{-1}$) to human fibrinogen (7.5 mg ml $^{-1}$, 0.15 M NaCl) in the presence of 10 mM CaCl $_2$ and 5 mM cysteine. The clot formed within 2 min. After 18 min, 0.05 original volumes of a trypsin solution (1 mg ml $^{-1}$) was added and the digestion allowed to proceed for 4 h at room temperature, at which point 0.05 original volumes of a soybean trypsin inhibitor solution (3 mg ml $^{-1}$) was added to stop further digestion. The solution was passed over a Sephadex G-150 gel filtration column (2.5 $\times$ 85 cm, 4 °C) that had been equilibrated with 1 M NaBr, 0.05 M sodium acetate, pH 5.3. The leading peak was pooled and precipitated by adding 0.33 volume saturated ammonium sulphate. The suspension was chilled on ice for 15 min before centrifugation, after which the supernatant was discarded and the pellet stored frozen at -20 °C until needed. Unreduced preparations gave essentially single bands with an apparent M_r of 170K on SDS gels. Precipitates were dissolved in minimal volumes of 0.05 M Tris buffer, pH 7.5, 0.005 M CaCl $_2$, quick-dialysed against that same buffer, adjusted to 10 mg ml $^{-1}$, and set for crystallization.

Crystals were obtained under similar conditions to those that were effective for native fragment D³³, except that the CaCl $_2$ concentration was reduced to 10 mM and the Gly-Pro-Arg-Pro-amide was set at 5 mM. Sitting drops were maintained at room temperature; the exact well conditions were: 13–14% PEG, 50 mM Tris, pH 7.0, 10 mM CaCl $_2$, 5 mM Gly-Pro-Arg-Pro-amide, 1 mM Na $_3$ N.

Crystallographic methods and data processing. Initial data collection and heavy-metal screening was conducted at room temperature on a Rigaku RU200 rotating anode generator and two area detectors (San Diego Multiwire Systems). Programs included with the San Diego Modelling System were used to scale and merge these data. The higher-resolution data used in the final model of fragment D were collected at the Daresbury synchrotron and Stanford Synchrotron Radiation Laboratory (Table 1). The double-D data were collected at the National Synchrotron Light Source, Brookhaven National Laboratory. Data collected on image plates were scaled with Denzo and ScalePack³⁴.

Fragment D crystals soaked in 1 mM KAuCl_4 for 6 days yielded reproducible gold derivatives as identified by peaks on the Harker section of a standard difference Patterson. The Harker section from the difference Patterson map contained four distinct peaks, but the maps could not be solved by conventional Patterson-solving programs or manually³⁵. Accordingly, six different data sets for gold derivatives were collected and the resulting isomorphous Patterson maps averaged together in Patterson space. This procedure enhanced the pattern on the Harker section and revealed that the four peaks were actually cross-peaks for two sets of atoms lying coplanar with each other at right angles to the crystallographic γ -axis. Two additional heavy-metal derivatives (uranium and mercury) were subsequently identified by difference Fourier techniques. The two-site second derivative was identified from crystals soaked in 1 mM $\text{UO}_2(\text{NO}_3)_2$ for 10 days, as well as the third derivative from crystals soaked in 1 mM HgI_4 ($\text{HgI}_2 + \text{excess KI}$) for 6 days.

Heavy-atom refinement and phasing was carried out with the PHASES package³⁶. The positioning of the heavy atoms and the poor quality of the initial solvent-flattened mask cast doubt on the exact boundaries of the molecular envelope. Great improvement was attained by a simple procedure of envelope-skeletonization with a program called Envder (G.S., manuscript in preparation). In brief, a search is made of the asymmetric unit (solvent envelope) for the largest sphere that can be contained within the envelope without penetrating solvent and not contained wholly within other spheres. A seed sphere is then set and other spheres connected to it by determining the mutual overlap. If the overlap exceeds a user-defined value, the spheres are considered to be connected. The operation is repeated until the number of monomers (two in this case) in the asymmetric unit is reached. The centre of mass of the monomers can be determined from this, and so, in combination with self rotation angles, can an approximate translation vector. The molecular envelope is then used to refine the n.c.s. matrices against the correlation coefficient.

The envelope was further improved by correlation mapping³⁷ and averaging and limited phase extension with the RAVE package³⁸ and DM³⁵. Model building was performed with O³⁹, and molecular replacement was done with X-PLOR⁴⁰, as were the structure refinements. SigmaA-weighting was performed as described⁴¹. The structure-homology search was performed with DALI⁴². Figures were created from the following programs: MOLSCRIPT^{43,44}; RASTER3D^{45,46}; GRASP⁴⁷; ALSCRIPT⁴⁸.

Received 20 June; accepted 8 August 1997.

1. Bailey, K., Astbury, W. T. & Rudall, K. M. Fibrinogen and fibrin as members of the keratin-myosin group. *Nature* **151**, 716–717 (1943).
2. Hall, C. E. & Slayter, H. S. The fibrinogen molecule: its size, shape and mode of polymerization. *J. Biophys. Biochem. Cytol.* **5**, 11–15 (1959).
3. Cohen, C. Invited discussion at 1960 Symposium on Protein Structure. *J. Polymer Sci.* **49**, 144–145 (1961).
4. Doolittle, R. F., Cassman, K. G., Cottrell, B. A., Friezner, S. J. & Takagi, T. Amino acid sequence studies on the α -chain of human fibrinogen. The covalent structure of the α -chain portion of fragment D. *Biochemistry* **16**, 1710–1710 (1977).
5. Laudano, A. P. & Doolittle, R. F. Studies on synthetic peptides that bind to fibrinogen and prevent fibrin polymerization. Structural requirements, numbers of binding sites and species differences. *Biochemistry* **19**, 1013–1019 (1980).
6. Chen, R. & Doolittle, R. F. γ - γ Crosslinking sites in human and bovine fibrin. *Biochemistry* **10**, 4486–4491 (1971).
7. Pizzo, S. V., Schwartz, M. L., Hill, R. L. & McKee, P. A. The effects of plasmin on the subunit structure of human fibrin. *J. Biol. Chem.* **248**, 4574–4583 (1973).
8. Hudry-Clergeon, G., Patural, L. & Suscillon, M. Identification d'un complexe (D-D) E dans les produits de dégradation de la fibrine bovine stabilisée par le facteur XIII. *Pathol. Biol.* **22** (suppl.), 47–52 (1974).
9. Rao, S. P. S. *et al.* Fibrinogen structure in projection at 18 Å resolution electron density by co-ordinated cryo-electron microscopy and X-ray crystallography. *J. Mol. Biol.* **222**, 89–98 (1992).
10. Yee, V. C. *et al.* Crystal structure of a 30 kDa C-terminal fragment from the γ chain of human fibrinogen. *Structure* **5**, 125–138 (1997).
11. Dang, C. V., Ebert, R. F. & Bell, W. R. Localization of a fibrinogen calcium binding site between γ -subunit positions 311 and 336 by terbium fluorescence. *J. Biol. Chem.* **260**, 9713–9717 (1985).
12. Townsend, R. R. *et al.* Carbohydrate structure of human fibrinogen. *J. Biol. Chem.* **257**, 9704–9710 (1982).
13. Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. Structure of the influenza haemagglutinin at the pH of membrane fusion. *Nature* **371**, 37–43 (1994).
14. Bouma, H. III, Takagi, T. & Doolittle, R. F. The arrangement of disulfide bonds in fragment D from human fibrinogen. *Thromb. Res.* **13**, 557–562 (1978).
15. Doolittle, R. F. A detailed consideration of a principal domain of vertebrate fibrinogen and its relatives. *Protein Sci.* **1**, 1563–1577 (1992).
16. Yamazumi, K. & Doolittle, R. F. Photoaffinity labeling of the primary fibrin polymerization

- site: Localization of the label to tyrosine γ -chain Tyr-363. *Proc. Natl Acad. Sci. USA* **89**, 2893–2896 (1992).
17. Gerloff, D. E., Cohen, F. E. & Benner, S. A. A predicted consensus structure for the C terminus of the β and γ chains of fibrinogen. *Protein Struct. Funct. Genet.* **27**, 279–289 (1997).
18. Okumura, N. *et al.* Fibrinogen Matsumoto I: a γ 364 Asp $\gg$ His (GAT to AT) substitution associated with defective fibrin polymerization. *Thromb. Haemost.* **75**, 887–891 (1996).
19. Ebert, R. F. (ed.) *1994 Index of Variant Human Fibrinogens* (CRC Press, Boca Raton, FL, 1994).
20. Janin, J., Miller, S. & Chothia, C. Surface, subunit interfaces and interior of oligomeric proteins. *J. Mol. Biol.* **204**, 155–164 (1988).
21. Mosesson, M. W. *et al.* The role of fibrinogen D domain intermolecular association sites in the polymerization of fibrin and fibrinogen Tokyo II (γ 275 Arg to Cys). *J. Clin. Invest.* **96**, 1053–1058 (1995).
22. Donahue, J. P., Patel, H., Anderson, W. E. & Hawiger, J. Three-dimensional structure of the platelet integrin recognition segment of the fibrinogen γ chain obtained by carrier protein-driven crystallization. *Proc. Natl Acad. Sci. USA* **92**, 12178–12182 (1994).
23. Weisel, J. W., Francis, C. W., Nagaswami, C. & Marder, V. J. Determination of the topology of factor XIIIa-induced fibrin γ -chain cross-links by electron microscopy of ligated fragments. *J. Biol. Chem.* **268**, 26618–26624 (1993).
24. Mosesson, M. W., Siebenlist, K. R., Hainfeld, J. F. & Wall, J. S. The covalent structure of factor XIIIa crosslinked fibrinogen fibrils. *J. Struct. Biol.* **115**, 88–101 (1995).
25. Mosesson, M. W., Siebenlist, K. R., Amrani, D. L. & DiOrio, J. P. Identification of covalently linked trimeric and tetrameric D domains in crosslinked fibrin. *Proc. Natl Acad. Sci. USA* **86**, 1113–1117 (1989).
26. Siebenlist, K. R., Meh, D. A. & Mosesson, M. W. Plasma factor XIII binds specifically to fibrinogen molecules containing γ' chains. *Biochemistry* **35**, 10448–10453 (1996).
27. Mosesson, M. W., Hainfeld, J., Wall, J. & Haschemeyer, R. H. Identification and mass analysis of human fibrinogen molecules and their domains by scanning transmission electron microscopy. *J. Mol. Biol.* **153**, 695–718 (1981).
28. Veklich, Y. I., Gorkun, O. V., Medved, L. V., Nieuwenhuizen, W. & Weisel, J. W. Carboxyl-terminal portions of the α chains of fibrinogen and fibrin. Localization by electron microscopy and the effects of isolated α C fragments on polymerization. *J. Biol. Chem.* **268**, 13577–13585 (1993).
29. Schiell, W. J. G., Adams, H. P. H. M., Voskuilen, M., Tesser, G. J. & Nieuwenhuizen, W. Structural requirements of position A α -157 in fibrinogen for the fibrin-induced rate enhancement of the activation of plasminogen by tissue-type plasminogen activator. *Biochem. J.* **276**, 655–659 (1991).
30. Nieuwenhuizen, W., Vermond, A., Voskuilen, M., Traas, D. W. & Verheijen, J. H. Identification of a site in fibrin(ogen) which is involved in the acceleration of plasminogen activation by tissue-type plasminogen activator. *Biochim. Biophys. Acta* **748**, 86–92 (1983).
31. Pratt, K. P. *et al.* The primary fibrin polymerization pocket: three-dimensional structure of a 30-kDa C-terminal γ chain fragment complexed with the peptide Gly-Pro-Arg-Pro. *Proc. Natl Acad. Sci. USA* **94**, 7176–7181 (1997).
32. Mihalyi, E. & Godfrey, J. E. Digestion of fibrinogen by trypsin. II. Characterization of the large fragment obtained. *Biochim. Biophys. Acta* **67**, 90–103 (1963).
33. Everse, S. J., Pelletier, H. & Doolittle, R. F. Crystallization of fragment D from human fibrinogen. *Protein Sci.* **4**, 1013–1016 (1995).
34. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
35. Collaborative Computing Project Number 4 The CCP4 suite: Programs for Protein Crystallography Version 3.1. *Acta Crystallogr. D* **50**, 760–763 (1994).
36. Furey, W. & Swaminathan, S. Phases 95: A program package for processing and analyzing diffraction data from macromolecules. *Methods Enzymol.* (in press).
37. Stein, P. E. *et al.* The crystal structure of pertussis toxin. *Structure* **2**, 45–57 (1994).
38. Kleywegt, G. J. & Jones, T. A. From first map to final model. Joint CCP4 and ESRF-EACBM. *Newslett. Protein Crystallogr.* **31**, 59–66 (1994).
39. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
40. Brunger, A. T. *X-PLOR. Version 3.1. A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, CT, 1992).
41. Read, R. J. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A* **42**, 140–149 (1986).
42. Holm, L. & Sander, C. Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138 (1993).
43. Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
44. Esnouf, R. M. An extensively modified version of Molscript which includes greatly enhanced colouring capabilities. *J. Mol. Graphics* **15**, 133–138 (1997).
45. Bacon, D. J. & Anderson, W. F. A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graphics* **6**, 219–220 (1988).
46. Merritt, E. A. & Murphy, M. E. P. Raster3D version 2.0: a program for photorealistic molecular graphics. *Acta Crystallogr. D* **50**, 869–873 (1994).
47. Nicholls, A., Sharp, K. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Protein Struct. Funct. Genet.* **11**, 281–296 (1991).
48. Barton, G. J. ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.* **6**, 37–40 (1993).
49. Pan, Y. & Doolittle, R. F. cDNA sequence of a second fibrinogen α chain in lamprey: an archetypal version alignable with full-length β and γ chains. *Proc. Natl Acad. Sci. USA* **89**, 2066–2070 (1992).
50. Fu, Y., Cao, Y., Hertzberg, K. M. & Grieninger, G. Fibrinogen α genes: conservation of bipartite transcripts and carboxy-terminal-extended α subunits in vertebrates. *Genomics* **30**, 71–76 (1995).

Acknowledgements. We thank J. Kraut for the use of his X-ray facilities; M. Sawaya for his time; H. Pelletier for help in establishing the initial crystallization conditions; D. Stuart for assistance; N. Xuong and J. Noel for access to equipment; R. Sweet at the Brookhaven National Laboratory for assistance; G. Walter for encouragement; and M. Riley and L. Veerapandian for technical assistance. This work was supported by an NIH grant and an American Heart Association postdoctoral fellowship to S.J.E.

Correspondence and requests for materials should be addressed to R.D. (e-mail:rdoolittle@ucsd.edu). Coordinates have been deposited in the Brookhaven Protein Data Bank (accession nos: fragment D, 1FZA; double-D, 1FZB).

Irreversibility of classical fluctuations studied in analogue electrical circuits

D. G. Luchinsky*† & P. V. E. McClintock*

* School of Physics and Chemistry, Lancaster University, Lancaster, LA1 4YB, UK

Fluctuations around some average or equilibrium state arise universally in physical systems. Large fluctuations—fluctuations that are much larger than average—occur only rarely but are responsible for many physical processes, such as nucleation in phase transitions, chemical reactions, mutations in DNA sequences, protein transport in cells and failure of electronic devices. They lie at the heart of many discussions^{1–5} of how the irreversible thermodynamic behaviour of bulk matter relates to the reversible (classical or quantum-mechanical) laws describing the constituent atoms and molecules. Large fluctuations can be described theoretically using hamiltonian^{6,7} and equivalent path-integral^{8–12} formulations, but these approaches remain largely untested experimentally, mainly because such fluctuations are rare and also because only recently was an appropriate statistical distribution function formulated¹¹. It was shown recently, however, that experiments on fluctuations using analogue electronic circuits allow the phase-space trajectories of fluctuations in a dynamical system to be observed directly¹². Here we show that this approach can be used to identify a fundamental distinction between two types of random motion: fluctuational motion, which takes the system away from a stable state, and relaxational motion back towards this state. We suggest that macroscopic irreversibility is related to temporal asymmetry of these two types of motion, which in turn implies a lack of detailed balance and corresponds to non-differentiability of the generalized non-equilibrium potential in which the motion takes place.

In a typical large fluctuation, a component x of some dynamical variable $\mathbf{x}$ departs temporarily from its stable state, moves to a remote state x_f and then returns: see, for example, Fig. 1a. The variable x might represent the voltage(s) in an electrical circuit (see below), or the phase of the order parameter in a superconducting quantum interference device¹³ (SQUID) or in an optical bistable device¹⁴, or the number densities of species in a chemical reaction¹⁵. Approximate theories of such phenomena have been developed^{6,7,9–12,15–21}, the approach being typically as follows. We consider overdamped brownian motion in a force field $\mathbf{K}(\mathbf{x}, t)$, assumed in general to be non-adiabatic and/or non-gradient (not derivable from a potential), driven by weak white noise $\xi(t)$ whose intensity $D \ll 1$ is considered to be the smallest parameter of the problem:

$$\dot{\mathbf{x}}_i = \mathbf{K}_i(\mathbf{x}, t) + \xi_i(t), \quad \langle \xi_i(t) \rangle = 0, \quad \langle \xi_i(t) \xi_j(0) \rangle = D \delta_{ij} \delta t \quad (1)$$

The corresponding Fokker–Planck equation for the probability density $P(\mathbf{x}, t)$ is:

$$\frac{\partial P(\mathbf{x}, t)}{\partial t} = -\nabla \cdot (\mathbf{K}(\mathbf{x}, t) P(\mathbf{x}, t) + \frac{D}{2} \nabla^2 P(\mathbf{x}, t)) \quad (2)$$

A large fluctuation (of scale $\gg \sqrt{D}$) consists of brownian motion away from a stable stationary state S . It is similar to quantum-mechanical tunnelling through a potential barrier, and can be treated in a similar way. In the limit of weak noise it can be

described by the Wentzel–Kramers–Brillouin (WKB or eikonal) approximation of the Fokker–Planck equation in the form $P(\mathbf{x}, t) = z(\mathbf{x}, t) \exp(W(\mathbf{x}, t)/D)$. Here $z(\mathbf{x}, t)$ is a prefactor, and $W(\mathbf{x}, t)$ is a classical action satisfying the Hamilton–Jacobi equation, which can be solved by integrating the hamiltonian equations of motion⁷:

$$\dot{\mathbf{x}} = \mathbf{p} + \mathbf{K}, \quad \dot{\mathbf{p}} = -\frac{\partial \mathbf{K}}{\partial \mathbf{x}} \mathbf{p}, \quad (3)$$

$$H(\mathbf{x}, \mathbf{p}, t) = \mathbf{p} \mathbf{K}(\mathbf{x}, t) + \frac{1}{2} \mathbf{p}^2, \quad \mathbf{p} \equiv \nabla W$$

where $\mathbf{p}$ is a momentum. We emphasise that equations (3) describe an auxiliary hamiltonian system which should be carefully distinguished from the microscopic hamiltonian equations of motion in a thermal system. We note also that the method is readily generalized, for example to systems driven by coloured noise^{10,20,21}.

The set of trajectories of equations (3) approaching the stable state S ($\bar{x} = x_s$, $\bar{p} = 0$) for positive time forms the stable invariant manifold of S , defined by the additional condition $\mathbf{p} = 0$. On this manifold, equations (3) reduce to $\dot{\mathbf{x}} = \mathbf{K}$, describing deterministic relaxation, so that these trajectories must correspond to the return path of the initial system (equations (1)) from x_f to S . The set of trajectories moving away from S in positive time forms the unstable invariant manifold of S , upon which $\mathbf{p} \neq 0$ in equations (3); they can be identified as describing the fluctuational paths from S to x_f . In each case, the trajectories represent^{6,7} optimal paths along which the system is expected to move with overwhelming probability during the fluctuation. But, although the decay of the fluctuation was well understood²², it remained uncertain for a long time whether or how the growth of a fluctuation along a trajectory of the unstable manifold might be observed experimentally. One problem was that noise in real systems is always of finite intensity, so that the relevance of the $D \rightarrow 0$ theory, for which the probability of having a large fluctuation at all also vanishes, was unclear. An additional puzzle originated in the analogy with optics and quantum mechanics²³ which, supplemented by a large body of numerical results, suggested^{15,18–21,24–26} that patterns of optimal fluctuational paths in general display singular features, but ones that differ significantly^{18,24,25} from those familiar from optics. It was sometimes inferred that the approximations leading to equations (3) would fail near the singularity points where, instead, a complete solution of the Fokker–Planck equation (equation (2)) would be required²⁷.

More recently, however, the physical reality both of the paths, and of the singular features in their pattern, have been demonstrated^{11,12} in analogue electronic experiments. We note that if the (approximate) hamiltonian description is correct, similar behaviour is to be expected of any fluctuating system describable in terms of equations (1); analogue electronic circuits are especially advantageous for such studies because their parameters are well controlled. The technique^{11,28} involves building an electronic circuit to model the system of interest, and driving it with external noise. Provided that no additional forces are applied, the system can be considered to be in thermal equilibrium at a temperature determined by the noise intensity and the damping constant, which are linked by the fluctuation dissipation theorem³. The state of the system is monitored continuously until eventually, as shown in Fig. 1a, a large fluctuation reaches the voltage x_f which we define to occur at $t = 0$. The interesting region of the path—including the fluctuational ($t < 0$) part f coming to x_f , as well as the relaxational ($t > 0$) part r leading back towards S —is then stored. An ensemble-average of such trajectories (Fig. 1b), built up over a period of time (typically weeks), creates the distribution $P_f(x, t) \equiv P(x_i, t_i; x, t; x_f, t_f)$, the probability of the system being at x at time t if it was at x_f at time t_f with $t_f = 0$ and $x_i = x_s$ at time $t_i = -\infty$. Unlike the original definition¹¹ of the prehistory distribution, $t > t_f$ is considered as well as $t < t_f$.

We now consider the application of these ideas to the fluctuations

† Permanent address: Russian Research Institute for Metrological Service, Ozernaya 46, 119361 Moscow, Russia.

of three very different example systems that can be used to describe a wide range of physical phenomena^{13–15,22,27,29}. The first (Fig. 1) is an overdamped Duffing oscillator, modelling a bistable system in thermal equilibrium near one of its stable states. Figure 1b shows a measured $P_f(x, t)$. We found: (1) that the relaxational and fluctuational parts of the distribution are symmetrical, which would only be expected⁴ under conditions of detailed balance (that is, when every transition between two states in one direction is on average balanced by a transition in the opposite direction); (2) that the ridges (modes) of the distribution follow closely the deterministic trajectory found from equations (3), plotted as full curve in Fig. 1c; and hence (3) that, in the macroscopic limit, where the width of the distribution tends to zero (because $D \rightarrow 0$) and one observes only the positions of the ridges, the paths to/from x_f themselves become reversible in time^{7,15,21,26,30}. Even in thermal equilibrium, however, hamiltonian theory envisages the fluctuational and relaxational trajectories as belonging to two different manifolds of the system (equations (3), with $\mathbf{p} \neq 0$ and $\mathbf{p} = 0$

respectively (see above). In the particular case of our analogue electronic model, this feature can be illustrated experimentally because the noise is external and a direct determination of p during fluctuations is therefore possible (compare the idea^{8–10,20} of the 'optimal force'). Thus we have been able to make simultaneous measurements of $x(t)$ and of the corresponding noise histories $\xi(t)$ causing transitions between the potential wells, setting $x_f = 0$ on the local potential maximum: some results are shown in Fig. 1d. Although the statistics of these very rare events are relatively poor, the data clearly demonstrate: first, that, as anticipated, $p \neq 0$ during the fluctuational part of the trajectory; second, that $p = 0$ within experimental error during relaxation; and third, that the hamiltonian theory (curves) describes very well both parts of the trajectory. Thus the time reversal symmetry can be regarded as arising from a degeneracy between the projections of two different curves in an extended (by the $\mathbf{p}$ -dimension(s)) phase space onto the space of the dynamical variables. Such an extension of the phase space seems an unnecessary complication in equilibrium; but it provides the key to

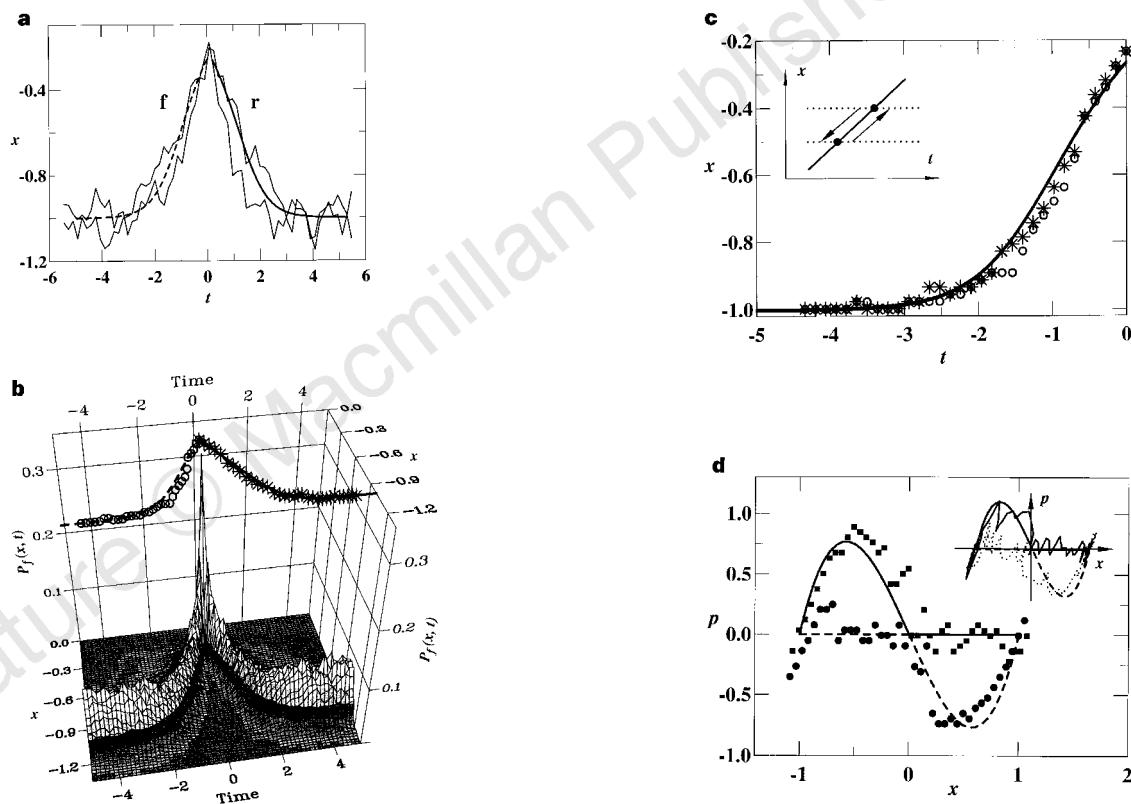


Figure 1 a–d, Fluctuational behaviour measured and calculated for an electronic model system in equilibrium: a double-well Duffing oscillator with $K(x) = x - x^3$, for $D = 0.014$. The model was constructed²⁸ from standard electronic components (operational amplifiers and analogue multipliers), tested for accuracy, and then treated as an entity in its own right—an object on which experiments could be performed. It was driven by external noise (a fluctuating voltage) from a noise generator, causing it to fluctuate about its equilibrium state at -1.0 V, representing $x = x_s = -1$ in equations (1). This response, a fluctuating voltage representing $x(t)$, was then digitized and analysed with a digital data processor. In particular, $x(t)$ was monitored continuously, waiting for large fluctuations reaching a chosen preset voltage threshold representing x_f . Fluctuational paths $x(t)$ reaching x_f were preserved for later analysis, and so also were the relaxational paths $x(t)$ from x_f back towards the stable state at $x = -1$. **a**, Two typical fluctuations (jagged lines) from the stable state at $S = -1$ to the remote state $x_f = -0.1$, and back again, are compared with the deterministic (noise-free) relaxational path from x_f to S (full, smooth, curve) and its time-reversed ($t \rightarrow -t$) mirror image (dashed curve). The fluctuational and relaxational parts of the trajectory are labelled f and r respectively.

In **b**, the lower plot gives the probability distribution $P_f(x, t)$ built up by ensemble-averaging a sequence of trajectories like those in **a**. The upper plot gives the positions of the ridges of $P_f(x, t)$ for the fluctuational (open circles) and relaxational (asterisks) parts of the trajectory for comparison with theoretical predictions (curves) based on equations (3). **c**, Plots of the ridges of $P_f(x, t)$, with $t \rightarrow -t$ for the relaxational part, demonstrating the time-reversal symmetry (main figure). The full curve is the theoretical prediction based on equations (3). Inset: detailed balance means that, for every transition in one direction between any two levels (for example, those shown by the dashed lines), there must be a return transition in the opposite direction; that is, given that a relaxational trajectory exists, the corresponding fluctuational one must be its time-inversed image in configuration space, so that detailed balance implies time-reversal symmetry. **d**, Demonstration of time-irreversible features of the fluctuations (main figure). The inset shows $p(x)$ for two typical transitional paths, from $x = -1$ to $x = 1$ (squares), and the reverse transition (filled circles). The full and dashed curves in the main figure and the inset are the corresponding paths predicted from equations (3).

understanding systems that are far from equilibrium, where the degeneracy is lifted by the presence of an external field.

Our second example (Fig. 2), is the archetypal nonequilibrium system considered by Graham^{16,17} in his discussion of the properties of the generalized potential, and further analysed in great detail by M. I. Dykman and V. N. Smelyanskiy (manuscript in preparation): an overdamped Duffing oscillator driven from equilibrium by a periodic force. The ridges of the fluctuational and relaxational parts

of the measured $P_f(x, t)$ are narrow and differ markedly in shape. The paths that they trace out (data points), compared in Fig. 2a with theoretical predictions (curves) calculated from equations (3), clearly demonstrate that the most probable fluctuational trajectory to (x_f, t_f) does *not* correspond to what one would obtain by time-reversing the relaxational trajectory. Figure 2b shows the measured $P_f(x, t)$ when x_f is placed on the calculated 'switching line'^{18,24}, a singularity separating regions that are approached via different

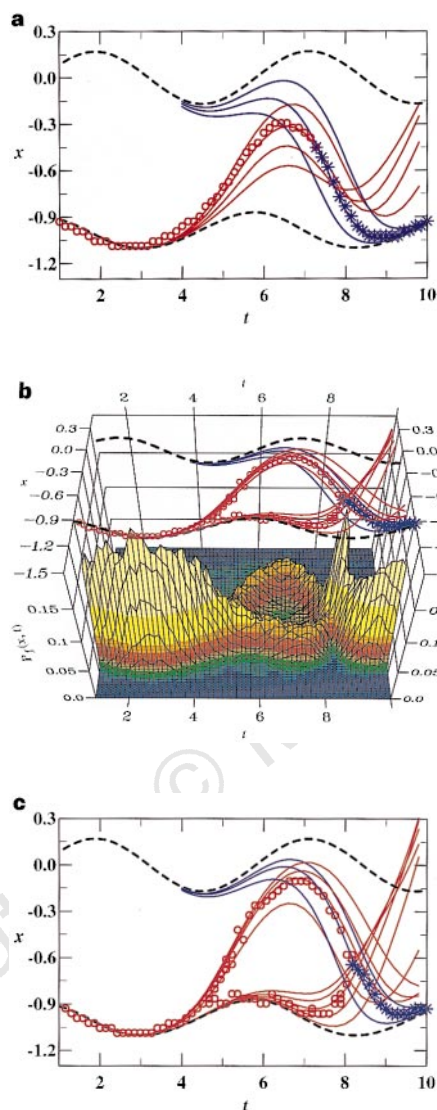


Figure 2 a–c, Fluctuational behaviour measured and calculated for an electronic model of a nonequilibrium system with explicit time dependence: an overdamped double-well Duffing oscillator with $K(x, t) = x - x^3 + A \cos(\omega t)$ and $A = 0.264$, $D = 0.012$. **a**, Fluctuational and relaxational paths (red circles and blue asterisks respectively) to/from the remote state $x_f = -0.46$, $t = 0.73$, found by tracing the ridges of a measured $P_f(x, t)$ distribution. The time-dependent stable and unstable states near $x = -1$ and $x = 0$ are shown by dashed lines. The fluctuational and relaxational paths calculated from equations (3) are shown as red and blue lines respectively. **b**, The measured $P_f(x, t)$ for a remote state $x_f = -0.63$, $t = 0.83$ that lies on the switching line. (lower plot), with contour colour shading to indicate constant levels of $P_f(x, t)$. In the upper plot fluctuational (red circles) and relaxational (blue asterisks) paths determined by tracing ridges of the distribution are compared with the corresponding (red and blue) theoretical lines predicted from equations (3). The dashed lines represent the time-dependent stable and unstable state, near $x = -1$ and $x = 0$ respectively. **c**, More detailed plot of the top-plane of **b**.

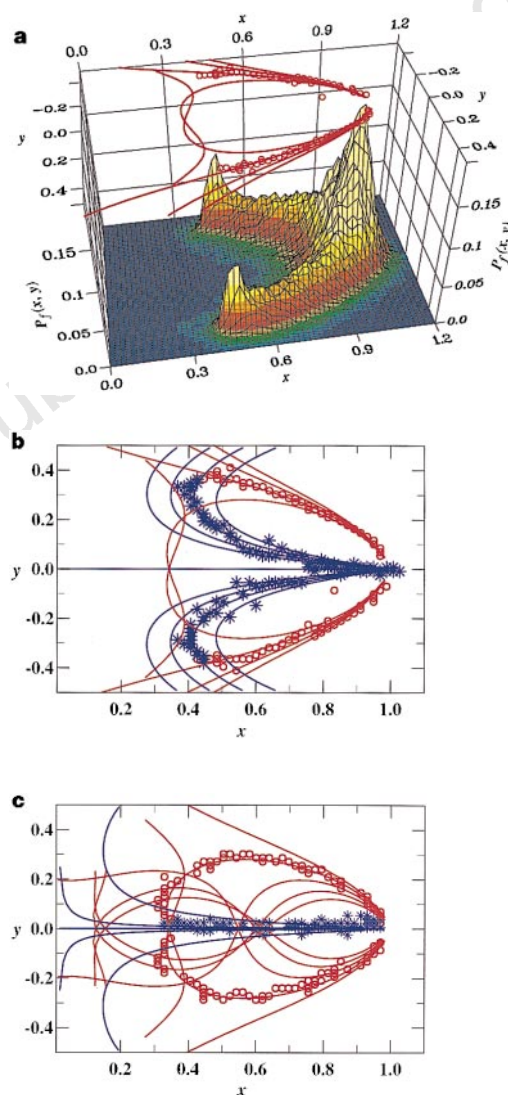


Figure 3 Fluctuational behaviour measured and calculated for an electronic model of a nonequilibrium system with a stationary nongradient field: $K(x, y) = (x - x^3 - axy^2, -(1 + x^2)y)$; $a = 10$; $D = 0.014$. **a**, The $P_f(x, y)$ distribution created by ensemble-averaging fluctuational paths leading from $S = (1, 0)$ to remote points at $\mathbf{x}_f = (0.44 \pm 0.35)$. (lower plot), with contour colour shading to indicate constant levels of $P_f(x, y)$. Plotted in the top-plane are the fluctuational paths to the remote state (red circles) found by tracing out the ridges of $P_f(x, y)$, compared to the predictions (red lines) of equations (3). **b**, Paths traced out by the ridges of the distribution in **a** for fluctuational motion (red circles), and by those of the corresponding distribution for relaxational motion (blue asterisks), compared with fluctuational (red) and relaxational (blue) optimal paths calculated from equations (3). **c**, as in **b**, but for the single remote state $\mathbf{x}_f = (0.32, 0)$ on the switching line.

fluctuational paths. The relaxational tail leading back to the stable state S is common to the two fluctuational paths that form the resultant corral¹². From Fig. 2c we see that the ridges of the distribution are strongly asymmetric in time, but agree well with the fluctuational and relaxational paths predicted from equations (3). The observed asymmetry of the distribution implies⁴ a lack of detailed balance; and it leads directly^{15,26} to macroscopic irreversibility in the $D \rightarrow 0$ limit where the width of the distribution tends to zero. In verifying the existence of the switching line, the results demonstrate the non-differentiability of the generalized nonequilibrium potential.

One third example (Fig. 3) is the system suggested by Maier and Stein^{19,26} for analysis of the escape problem in nonequilibrium systems. It consists of a bistable system driven from equilibrium by a stationary field of (in general) the nongradient type. In measurements near one of the stable states, the lack of detailed balance and the irreversibility of the fluctuations become strikingly apparent. Figure 3a shows the measured $P_f(x, y)$ distribution for fluctuations to the two different remote states x_f placed symmetrically on either side of the y axis; $P_f(x, y)$ is the projection of $P_f(x_i, y_i, t_i; x_f, y_f, t_f)$ onto the x - y plane with $t_i = 0$ as before. When the paths traced out by the ridges are plotted (Fig. 3b), it can be seen: (1) that the fluctuational trajectories are completely different from the relaxational ones; and (2) that they are in good agreement with the fluctuational and relaxational trajectories predicted from equations (3). Figure 3c shows the corresponding picture measured for fluctuations to a single remote state x_f on the switching line (lying on the x axis). The two fluctuational paths to the remote state are, again, markedly different from the common relaxational path leading back to S . Both parts of the trajectory are well described by the corresponding optimal paths calculated from equations (3). In this case too, therefore, the observations link a lack of detailed balance to macroscopic irreversibility, and demonstrate the nondifferentiability of the generalized nonequilibrium potential; the closed loops traversed during fluctuations verify the expected^{2,26,27} occurrence of rotational flow in nonequilibrium systems.

Thus, after more than 60 years of "thought experiments" on large fluctuations^{2,3,30}, it has now become possible to do real experiments yielding quantitative results. Our work has already verified several long-standing theoretical predictions, including those of symmetry between the growth and decay of classical fluctuations in equilibrium^{2,30}, the breaking of this symmetry under nonequilibrium conditions^{7,15,26}, the relationship of symmetry-breaking to a lack of detailed balance^{15-17,26}, and the existence of an optimal force derived from the fluctuating field and related to the momentum of an auxiliary hamiltonian system^{8-10,20,21}. The technique has also enabled us to reveal new dynamical features of large fluctuations, such as critical broadening of the prehistory probability distribution¹². It will be as applicable to future experiments on natural systems as it has been to the electronic models studied here. Because equations (1) can, at least in some cases^{9,31}, be related directly to the microscopic classical equations of motion for a system interacting with a heat bath, the approach illuminates connections between microscopic reversibility and macroscopic irreversibility. Other open questions needing urgently to be addressed include the precise physical meaning of the momentum p and the role of entropy and how it changes during fluctuational motion. □

Received 7 January; accepted 23 July 1997.

1. Einstein, A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys.* **17**, 549–560 (1905); also transl. in *Investigations on the Theory of the Brownian Movement* (ed. Fürth, R.) (Dover, New York, 1956).
2. Onsager, L. Reciprocal relations in irreversible processes. I. *Phys. Rev.* **37**, 405–426 (1931).
3. Landau, L. D. & Lifshitz, E. M. *Statistical Physics Part I*, 3rd edn (Pergamon, New York, 1980).
4. Schulman, L. S. Models for intermediate time dynamics with two-time boundary conditions. *Physica A* **117**, 373–380 (1991).
5. Bricmont, J. Science of chaos or chaos in science? *Ann. NY Acad. Sci.* **775**, 131–175 (1996).

6. Cohen, J. K. & Lewis, R. M. A ray method for the asymptotic solution of the diffusion equation. *J. Inst. Math. Appl.* **3**, 266–290 (1967).
7. Freidlin, M. I. & Wentzell, A. D. *Random Perturbations in Dynamical Systems* (Springer, New York, 1984).
8. Feynman, R. P. & Hibbs, A. R. *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
9. Dykman, M. I. & Krivoglaz, M. A. in *Soviet Physics Reviews* Vol. 5 (ed. Khalatnikov, I. M.) 265–441 (Harwood, New York, 1984).
10. Bray, A. J. & McKane, A. J. Instanton calculation of the escape rate for activation over a potential barrier driven by colored noise. *Phys. Rev. Lett.* **62**, 493–496 (1989).
11. Dykman, M. I., McLintock, P. V. E., Smelyanskiy, V. N., Stein, N. D. & Stocks, N. G. Optimal paths and the prehistory problem for large fluctuations in noise-driven systems. *Phys. Rev. Lett.* **68**, 2718–2721 (1992).
12. Dykman, M. I., Luchinsky, D. G., McLintock, P. V. E. & Smelyanskiy, V. N. Coralls and critical behaviour of the distribution of optimal paths. *Phys. Rev. Lett.* **77**, 5229–5232 (1996).
13. Kautz, R. L. Noise, chaos, and the Josephson standard. *Rep. Prog. Phys.* **59**, 935–992 (1996).
14. Dykman, M. I. et al. Noise-enhanced optical heterodyning. *Appl. Phys. Lett.* **67**, 308–310 (1995).
15. Dykman, M. I., Mori, E., Ross, J. & Hunt, P. M. Large fluctuations and optimal paths in chemical kinetics. *J. Chem. Phys.* **100**, 5737–5750 (1994).
16. Graham, R. & Tél, T. Existence of a potential for dissipative dynamical systems. *Phys. Rev. Lett.* **52**, 9–12 (1984).
17. Graham, R. in *Noise in Nonlinear Dynamical Systems* Vol. 1 (eds Moss, F. & McLintock, P. V. E.) 225–278 (Cambridge Univ. Press, 1989).
18. Dykman, M. I., Millonas, M. M. & Smelyanskiy, V. N. Observable and hidden features of large fluctuations in nonequilibrium systems. *Phys. Lett. A* **195**, 53–58 (1994).
19. Maier, R. S. & Stein, D. L. A scaling theory of bifurcations in the symmetrical weak-noise escape problem. *J. Statist. Phys.* **83**, 291–357 (1996).
20. Dykman, M. I. Large fluctuations and fluctuational transitions in systems driven by colored Gaussian noise—a high-frequency noise. *Phys. Rev. A* **42**, 2020–2029 (1990).
21. Einchcomb, S. J. B. & McKane, A. J. Use of Hamiltonian mechanics in systems driven by colored noise. *Phys. Rev. E* **51**, 2974–2981 (1995).
22. Van Kampen, N. G. *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 1990).
23. Berry, M. V. Waves and Thom's theorem. *Adv. Phys.* **25**, 1–26 (1976).
24. Jauslin, H. R. Nondifferentiable potentials for nonequilibrium steady states. *Physica A* **144**, 179–191 (1987).
25. Day, M. V. Recent progress on the small parameter exit problem. *Schrochastics* **20**, 121–150 (1987).
26. Maier, R. S. & Stein, D. L. Escape problem for irreversible systems. *Phys. Rev. E* **48**, 931–938 (1993).
27. Risken, H. *The Fokker-Planck Equation* 2nd edn (Springer, Berlin, 1989).
28. McLintock, P. V. E. & Moss, F. in *Noise in Nonlinear Dynamical Systems* Vol. 3 (eds Moss, F. & McLintock, P. V. E.) 243–274 (Cambridge Univ. Press, 1989).
29. Haken, H. Cooperative phenomena in systems far from thermal equilibrium and in nonphysical systems. *Rev. Mod. Phys.* **47**, 67–119 (1975).
30. Onsager, L. & Machlup, S. Fluctuations and irreversible processes. *Phys. Rev.* **91**, 1505–1512 (1953).
31. Zwanzig, R. Nonlinear generalised Langevin equations. *J. Statist. Phys.* **9**, 215–220 (1973).

Acknowledgements. We thank M. I. Dykman for help and encouragement; we also thank him, R. S. Maier and D. L. Stein for comments on an earlier version of the manuscript; and we acknowledge correspondence with N. G. van Kampen. This work was supported by the Engineering and Physical Sciences Research Council (UK), the Royal Society, and the Russian Foundation for Basic Research.

Correspondence should be addressed to P.V.E.M.C. (e-mail: p.v.e.mclintock@lancaster.ac.uk).

Light amplification in organic thin films using cascade energy transfer

M. Berggren, A. Dodabalapur, R. E. Slusher & Z. Bao

Bell Laboratories, Lucent Technologies, 600 Mountain Avenue, Murray Hill, New Jersey 07974, USA

There is currently renewed interest in the development of lasers using solid-state organic and polymeric materials as the gain media. These materials have a number of properties that make them good candidates for such applications—for example, emission bands that are displaced (via a Stokes shift) from absorption bands, and the ease with which the emitting species can be embedded in a suitable host material^{1–5}. But despite these advantages, the threshold power densities required for light amplification that have been reported so far have been high^{6–8}. Here we describe an approach, based on energy transfer between molecular species, that can lower the threshold for stimulated emission and laser action while improving markedly the waveguiding properties of the active material. In our materials, an initial molecular excited state is generated in the host compound by absorption of light; this state is then resonantly and non-radiatively transferred down in energy (through one or more steps)

between suitably matched dye molecules dispersed in the host, so ensuring that the absorption losses at the final emission wavelengths are very small. Such composite gain media provide broad tunability of the emission wavelength, and also decouple the optical emission properties from the transport properties, so providing greater flexibility for the design of future electrically driven device structures.

Solid-state lasers based on organic materials, with a few recent exceptions, have had gain media consisting of a suitable transparent host doped with a small percentage of dye^{2–5}. In such media there is no significant optical interaction between the host molecules and the dye molecules; the host serves mainly to physically separate the fluorescent dye molecules, which has the beneficial effect of hindering concentration quenching¹, and increasing the quantum yield of the material. As the host matrix is non-absorbing, and the volume fraction of the emissive dye is small (of the order of 1%), the amount of incident light that can be absorbed in a film of thickness ~150–200 nm (a typical value for the thickness that supports only the lowest-order optical mode) is very small (<5%). This places a lower limit on the threshold power density necessary to have stimulated emission.

One way to circumvent this limit is to have a host material which has optical gain. There have been several recent examples of such luminescence in conjugated polymers^{6–8}. It is also possible to obtain stimulated emission in pure films of small molecules, an example of which will be given below. Although the absorption of incident light can be high in such materials, the absorption losses at the emission wavelengths can be quite high (>1 cm⁻¹). This is because the absorption coefficient is quite high in most organic materials close to the absorption 'edge', and it is necessary to shift the luminescence (sometimes by >200 nm) from this edge to significantly lower the absorbance. A low absorbance at the emission wavelength(s) will mean that the optical gain and minimum pump power necessary for population inversion is lowered. One way to accomplish this is to use gain media with energy transfer^{9,10}. In the most general case, the energy transfer can take place through multiple steps, a process we call cascade energy transfer. This process requires three or more materials which have absorption and emission spectra that suitably matched. We believe that this report represents the first time cascade energy transfer has been used in an optical device. Single-step energy transfer can also be useful in laser gain media, and we give some examples of this later.

In the materials studied here, light is absorbed by host material

such as 2-(4-biphenyl)-5-(4-*t*-butylphenyl)-1,3,4-oxadiazole (PBD) which is doped with a small percentage (of the order of 1 wt%) of one or more fluorescent dyes with different absorption and emission spectra. The function of the host is to absorb the pump light and funnel the excitation to the dye molecules through a Förster-type energy transfer. Because the average spacing between the dye molecules at the doping concentrations we use (3 nm) is less than the Förster distance¹⁰, excitation transfer can be extremely efficient, especially if the overlap between the donor (host) emission spectrum and acceptor (dye) absorption spectrum is good. The absorption and emission spectra of the materials used are shown in Fig. 1. The spectra for PBD are from a pure film whereas the other spectra have been obtained from solid solutions of the dye in polystyrene. Efficient transfer was observed between PBD and coumarin 490 (C490) and between PBD and DCMII (with C490 included) and between PBD and LDS821 (with both C490 and DCMII also included in the film). (The dyes DCMII and LDS821 were obtained from Exciton, Inc., Dayton, Ohio). The transfer was poor between PBD and LDS821 when the intermediate dyes—C490 and DCMII—were not included in the film. These results are consistent with the predictions of resonance energy transfer theory^{9,10}. To minimize the effects of reabsorption and almost eliminate waveguiding, the films were very thin and were deposited on aluminium. In other experiments, the relative quantum yield of films of the various materials were estimated relative to the quantum yield of a neat film of PBD (taken as 1). The quantum yields measured are 1.67 (PCD + C490), 1.49 (PCD + C490 + DCMII) and 0.6 (PBD + DCMII + LDS821).

The films that we used to study stimulated emission were deposited on glass substrates. The host material in the first set of studies was PBD which, together with any dopants, was deposited by spin-coating to form ~200-nm-thick films. The thickness of the films is such that only the lowest-order mode is guided by the planar waveguide formed by the PBD (core) and air and glass (cladding) layers. These films were photopumped with unfocused light from a pulsed nitrogen laser (wavelength $\lambda = 337$ nm) which emits 2-ns-long pulses with an energy density of 500 $\mu\text{J cm}^{-2}$ (corresponding to 250 kW cm⁻²). The spot size was 3 mm $\times$ 2.5 cm. Neutral-density filters were used to adjust the pump power level. At low pump powers, the emission spectra are identical to the spontaneous emission spectra shown in Fig. 1. At sufficiently high pump powers the emission spectrum becomes narrower; such a narrowing is characteristic of amplification of spontaneous emission (ASE)

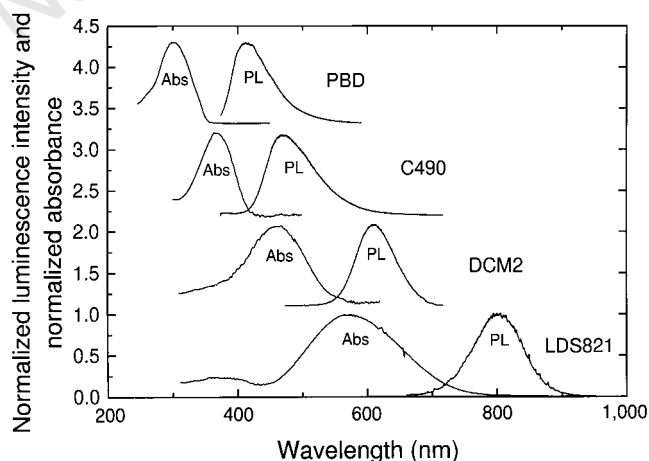


Figure 1 Absorbance and photoluminescence spectra of the host material and the fluorescent dyes. The host material is 2-(4-biphenyl)-5-(4-*t*-butylphenyl)-1,3,4-oxadiazole (PBD); the dyes are coumarin 490 (C490), DCMII and LDS821. The spectra for PBD are from a pure film whereas the other spectra are from solid solutions of the material in polystyrene.

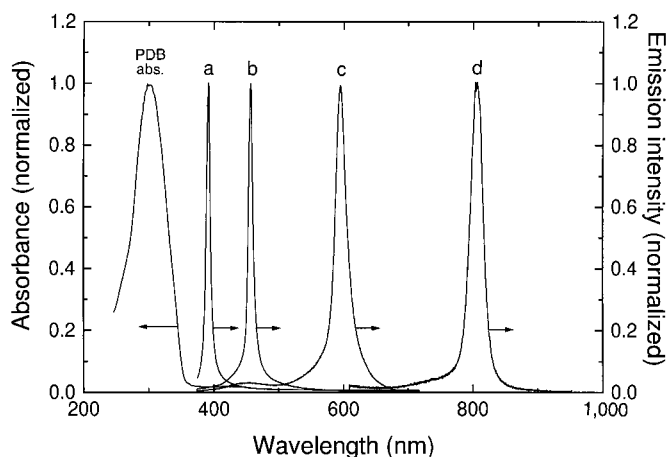


Figure 2 Normalized absorbance of PBD, and stimulated emission spectra from this material alone, and when doped with dyes. Curve a, PBD only; curve b, 0.6% C490 in PBD; curve c, 0.6% C490 + 0.9% DCMII in PBD; curve d, 0.6% C490 + 0.9% DCMII + 0.6% LDS821 in PBD.

and is observed in a wide variety of gain media. Films of undoped PBD also have such amplifying characteristics, as shown in Fig. 2 curve a. The threshold pump power (with a gain area of $3\text{ mm} \times 2.5\text{ cm}$) necessary to detect significant stimulated emission is $\sim 10\text{ kW cm}^{-2}$. One reason for this high value is that the peak of the ASE is close to the absorption 'edge' of the PBD where the absorption coefficient is high. As the gain coefficient has to be more than the absorption coefficient to have amplification, it follows that the pump power is necessarily high. Also shown in Fig. 2 are the ASE spectra (when pumped at 250 kW cm^{-2}) of PBD doped with 0.6% C490 (curve b) and 0.6% C490 + 0.9% DCMII (curve c). The peak of curve b which is at $\sim 460\text{ nm}$ coincides almost exactly with the peak of the C490 spontaneous emission spectrum. The spontaneous emission from PBD at 460 nm is quite weak and separate measurements showed that energy transfer from PBD to C490 is very efficient. Thus, we attribute the ASE peak at 460 nm as due primarily to amplification of spontaneous emission from C490 following energy transfer from PBD. Similarly, in curve c the ASE peak is almost exclusively from DCMII as both PBD and C490 have very little or no spontaneous emission intensities at wavelengths near 600 nm . We found that energy transfer between PBD and DCMII was not as efficient as between PBD and C490 or between C490 and DCMII. Cascade energy transfer from PBD to DCMII via C490 molecules is the most likely pathway for excitation transfer. When the third dye (LDS821) was also included in the films, ASE centred

near 805 nm was observed (curve d). The excitation transfer from PBD to LDS821 must proceed through one or more intermediate steps, as direct transfer from PBD to LDS821 is extremely inefficient. Based on studies of spontaneous emission in such doped films, we feel that the most likely pathway of excitation transfer is through both C490 and DCMII although two step transfers (through either C490 or DCMII) are also possible.

Figure 2 shows that pumping at 337 nm has resulted in a series of ASE peaks from the ultraviolet (392 nm) to the near infrared (805 nm). The thresholds for stimulated emission are the lowest for cases b, c and d where powers of $0.8\text{--}1.2\text{ kW cm}^{-2}$ are needed as opposed to 10 kW cm^{-2} for the pure film. As mentioned above, the chief reason for this is that the absorption coefficient is relatively low at wavelengths corresponding to peaks b, c and d.

We have fabricated distributed Bragg reflector lasers by pre-patterning a thermally oxidized Si wafer so as to form two gratings separated by distances in the range 3 mm to 1 cm . The gratings act as wavelength-selective mirrors. The grating period is $0.6\text{ }\mu\text{m}$ which results in reflectivity peaks near 610 nm , 485 nm and 392 nm , corresponding to the third, fourth and fifth orders, respectively. Spin-coating the active material between the two gratings completes the laser. The peaks in reflectivity of such gratings are matched to the fluorescence spectra of PBD, PBD doped with C490, and PBD doped with both C490 and DCMII, as evident from Fig. 2. We have used these active gain media together with the distributed Bragg reflector resonator to fabricate ultraviolet, blue and red lasers, as shown in Fig. 3.

The measurements reported above were all made with spun films. The film quality of vacuum-sublimed films is sometimes better than that of spun films. We now describe the characteristics of optical amplifiers in which the active materials is 2-naphthyl-4,5-bis(4-methoxyphenyl)-1,3-oxazole (NAPOXA) doped with $\sim 1\%$ DCMII. These films are deposited by co-sublimation and combine two important properties: good film-forming ability and almost perfect overlap between the host emission and guest absorption spectra, as can be observed in Fig. 4. With this system thresholds for stimulated emission were as low as 85 W cm^{-2} (total power). The pump energy density for these measurements was held constant at $0.1\text{--}0.2\text{ }\mu\text{J cm}^{-2}$, and the length of the gain region was varied in the range $1\text{--}3\text{ cm}$ while the width of the gain region was $3\text{--}4\text{ mm}$. We believe that this is the lowest threshold power density for an organic film reported to date and is due principally to the fact that the combined absorption and other losses of the NAPOXA host and the

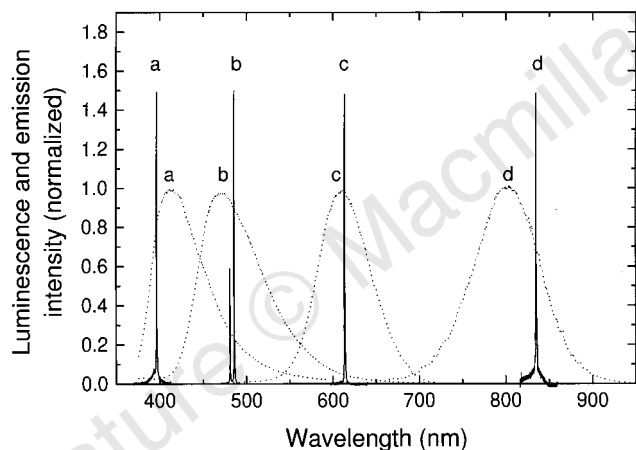


Figure 3 Laser emission spectra from devices made from four different active materials, and the spontaneous emission spectra of these materials. Solid lines, laser emission spectra; curve a, PBD; curve b, PBD + C490; curve c, PBD + C490 + DCMII; curve d, PBD + C490 + DCMII + LDS821. Dotted lines, spontaneous emission spectra for these active materials. The lasers producing emission spectra a, b and c used the same distributed Bragg reflector (DBR) resonator consisting of two $0.6\text{-}\mu\text{m}$ -pitch gratings etched in SiO_2 separated by a gain region $3\text{--}10\text{ mm}$ long. The third, fourth and fifth harmonics of the reflectivity of such DBR mirrors match the spontaneous emission spectra of c, b and a, respectively. The lasing threshold for pure films of PBD is quite high (10 kW cm^{-2}). The measured threshold power densities for the blue (b) and red (c) lasers are $1\text{--}1.5\text{ kW cm}^{-2}$ (total incident power density). The thresholds for the blue and red lasers are affected by losses in the SiO_2 cladding layer. As a result of these losses, the thresholds for observing ASE on these oxide-coated Si substrates (without a DBR) is $\sim 10\text{ kW cm}^{-2}$, much higher than the threshold for films on glass substrates. The laser thresholds can be further lowered by reducing optical losses in the SiO_2 . For the laser d, a two-dimensional photonic bandgap based distributed feedback geometry is used. This is realized by using a SiO_2 -coated Si substrate in which a square lattice of holes (with a circular cross-section and radius $0.15\text{ }\mu\text{m}$) were etched. The depth of the holes is $10\text{--}30\text{ nm}$ and the lattice constant is $0.6\text{ }\mu\text{m}$. The active material was spin-coated on such a patterned substrate to complete the device. The linewidths of the lasers a-d are resolution limited and are $0.2\text{--}0.3\text{ nm}$.

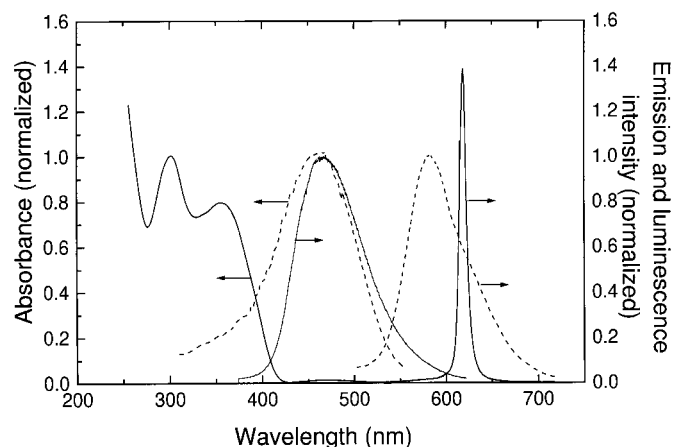


Figure 4 Absorbance and spontaneous emission spectra of the compounds used to make vacuum-sublimed films. Solid lines, 2-naphthyl-4,5-bis(4-methoxyphenyl)-1,3-oxazole (NAPOXA); dashed lines, DCMII. Also shown is the stimulated emission spectrum from a film of NAPOXA doped with $\sim 1\%$ DCMII when excited at 250 kW cm^{-2} .

dye dopants at amplification wavelengths is small ($<1\text{ cm}^{-1}$). The threshold power densities achieved ($<100\text{ W cm}^{-2}$) compare very favourably with those reported for unpatterned films of conjugated polymers^{6–8}. The low threshold powers in our system are a direct consequence of low loss at the emission wavelengths which allows the gain length to be increased to $>2\text{ cm}$.

Resonance energy transfer is fundamentally an electromagnetic interaction and the transfer rates can be influenced by strong microcavity effects. In high- Q cavities, this interaction between a donor and an acceptor is to be expected to be stronger. This means that it may be possible to use an even smaller doping concentration and still achieve nearly complete transfer. This would have the potential to further lower absorption losses and thresholds and is a promising area for research.

The beneficial effects of energy transfer are not restricted to small organic molecules but also apply in the case of longer molecules such as oligomers and polymers. We have obtained low threshold power densities (for ASE detection) of $\sim 200\text{ W cm}^{-2}$ for both polymer and oligomer emitters. The host material used is PBD and the emitters are a soluble derivative of poly(phenylene vinylene) (10 wt%) and an oligomeric phenylene-vinylene (2 wt%). We measured a combined waveguide loss (due to effects such as absorption and scattering) of 0.6 cm^{-1} in the case of the conjugated polymer emitter. Such low losses indicate that it is possible to use such materials in resonators with $Q > 10^5$, thereby accessing a regime where cavity electrodynamic effects can help to lower laser thresholds significantly. These results show the generality of our assertions that the absorption loss and threshold for stimulated emission is lowered, and the waveguiding properties improved, when the excitation is transferred down in energy in organic materials. □

Received 20 February; accepted 8 August 1997.

1. Schafer, F. P. in *Topics in Applied Physics: Dye Lasers* 3rd edn (ed. Schafer, F. P.) Ch. 1 (Springer, Berlin, 1990).
2. Weber, H. P. & Ulrich, R. A thin film ring laser. *Appl. Phys. Lett.* **19**, 38–40 (1971).
3. Kogelnik, H. & Shank, C. V. Stimulated emission in a periodic structure. *Appl. Phys. Lett.* **18**, 152–154 (1971).
4. Kaminov, I. P., Weber, H. P. & Chandross, E. A. Poly(methyl methacrylate) dye laser with internal diffraction grating. *Appl. Phys. Lett.* **18**, 497–499 (1971).
5. Gonokami, M. et al. Polymer microdisc and microring lasers. *Opt. Lett.* **20**, 2093–2095 (1995).
6. Tessler, N., Denton, G. J. & Friend, R. H. Lasing from conjugated polymer microcavities. *Nature* **382**, 695–697 (1996).
7. Hide, F. et al. Semiconducting polymers: a new class of solid state laser materials. *Science* **273**, 1833–1836 (1996).
8. Frolov, S. V. et al. Cooperative emission in π -conjugated polymer thin films. *Phys. Rev. Lett.* **78**, 729–736 (1997).
9. Förster, T. Transfer mechanisms of electronic excitation. *Disc. Faraday Soc.* **27**, 7–17 (1959).
10. Van Der Meer, B. W., Coker, G. & Chen, S. Y. in *Resonance Energy Transfer* Ch. 2 (VCH, New York, 1994).

Acknowledgements. We thank A. Timko and O. Nalamasu for the gratings, and E. A. Chandross and H. E. Katz for comments.

Correspondence and requests for materials should be addressed to A.D. (e-mail: ananth@bell-labs.com).

A chiral spherical molecular assembly held together by 60 hydrogen bonds

Leonard R. MacGillivray and Jerry L. Atwood

Department of Chemistry, University of Missouri-Columbia, Columbia, Missouri 65211, USA

Spontaneous self-assembly processes that lead to discrete spherical molecular structures are common in nature. Spherical viruses¹ (such as hepatitis B) and fullerenes² are well-known examples in which non-covalent and covalent forces, respectively, direct the assembly of smaller subunits into larger superstructures. A common feature of these shell-like architectures is their

ability to encapsulate neutral and/or charged guests whose size, shape and chemical exteriors complement those of the host's inner surface^{3,4}. Their interiors can often be regarded as a new phase of matter⁵, capable of controlling the flow of reactants, transients and products, and of catalysing reactions of both chemical and biological relevance. Such properties have inspired the recent emergence of monomolecular^{5–7} and supramolecular dimeric molecular capsules^{8,9}, many of which have been based on the head-to-head alignment of bowl-shaped polyaromatic macrocycles such as calix[4]arenes^{5,7,9}. But true structural mimicry of frameworks akin to viruses and fullerenes, which are based on the self-assembly of $n > 3$ subunits, and where surface curvature is supplied by edge sharing of regular polygons, has remained elusive. Here we present an example of such a system: a chiral spherical molecular assembly held together by 60 hydrogen bonds (1) (Fig. 1). We demonstrate the ability of 1, which consists of six calix[4]resorcinarenes 2 and eight water molecules, to self-assemble and maintain its structure in apolar media and to encapsulate guest species within a well-defined cavity that possesses an internal volume of about $1,375\text{ Å}^3$. Single crystal X-ray analysis shows that its topology resembles that of a spherical virus¹ and conforms to the structure of a snub cube, one of the 13 Archimedean solids¹⁰.

Our interest in 2 lies in its ability to function as a multiple hydrogen bond donor in the solid state¹¹. During experiments aimed at co-crystallizing C-methylcalix[4]resorcinarene 2a with hydrogen bond acceptors in aromatic solvents, we have discovered the ability of 2a to self-assemble as a spherical hexamer, along with adventitious water molecules, to form 1a (Fig. 1). Subsequent solution studies have revealed the ability of C-undecylcalix[4]resorcinarene 2b to maintain the structure of the spheroid in apolar organic solvents. We now report the synthesis, X-ray crystal structure and solution behaviour of 1.

Addition of 2a (0.015 g) to a boiling aliquot of nitrobenzene (5 ml) yielded light yellow cubic crystals 3 suitable for X-ray analysis after a period of ~ 2 weeks. The formulation of 3 was confirmed by single-crystal X-ray diffraction and ¹H NMR spectroscopy (see Supplementary Information).

A cross-sectional view of the X-ray crystal structure of 1a is shown in Fig. 1a. The assembly consists of six molecules of 2a and eight molecules of water that have assembled, by way of 60 O–H...O hydrogen bonds, to form a shell-like octahedral cubic spheroid. The calixarenes, each of which lies around a four-fold axis, point their hydroxy groups along the periphery of 1a and form two hydrogen bonds to two neighbouring calixarenes (O1...O1' 2.73(1) Å) whereas the water molecules, each of which lies around a three-fold axis, are embedded along the surface of 1a such that they lie on the vertices of a cube (edge length 9.00 Å) and participate in three hydrogen bonds with three different calixarenes (O2...O3 2.73(1) Å). Notably, each calixarene also exhibits four intramolecular hydrogen bonds, one at each of its corners (O1...O2 2.64(1) Å), which impart stability to its bowl-like conformation. As a result, 1a possesses 4-3-2 symmetry (Fig. 1b–d), ignoring all hydroxy hydrogen atoms, and a well-defined central cavity with a maximum diameter of 17.7 Å and an internal volume of $\sim 1,375\text{ Å}^3$ (Fig. 1e)¹². Indeed, the cavity of 1a is vast, being more than 4.5 times larger than the cavity of the largest molecular capsule reported to date ($\sim 300\text{ Å}^3$)⁸. Interestingly, the calixarenes of 1a are twisted by 23° with respect to the faces of the water 'cuboid' which, as a consequence, makes 1a chiral.

A view depicting the solid-state packing of 1a reveals that neighbouring spheroids fall on their four-fold axes and their methyl groups lie staggered owing to the twisting displayed by the macrocycles (Fig. 1f). This gives rise to an interpenetrating body-centred cubic lattice that exhibits interstices between 1a that are occupied by disordered water molecules, each of which participates in two hydrogen bonds with two water molecules from adjacent

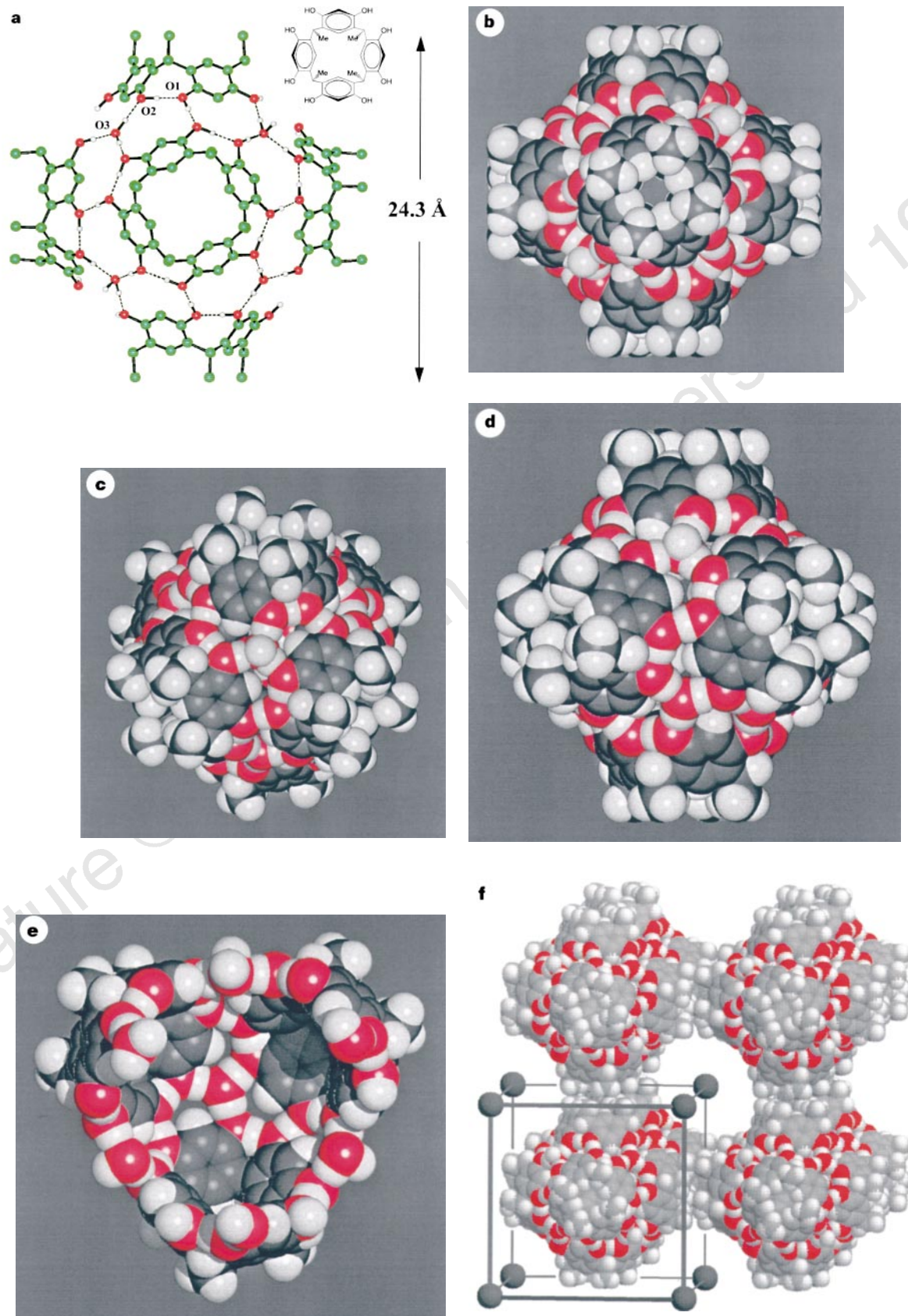


Figure 1 The structure of **1a**: **a**, cross-sectional view (inset, structural formula for **2a**); **b-d**, space-filling views along the crystallographic four-fold rotation axis (**b**), three-fold rotation axis (**c**), and two-fold rotation axis (**d**); **e**, cut-away view along the three-fold rotation axis; **f**, solid-state packing, where the shaded grey spheres also represent **1a** (solvent molecules have been omitted for clarity). The

spheroid is held together by 60 hydrogen bonds (hydrogen bond donor/hydrogen bond acceptor/hydrogen bond type/number of hydrogen bonds): calixarene/calixarene/intramolecular/24; calixarene/calixarene/intermolecular/12; calixarene/water/intermolecular/12; water/calixarene/intermolecular/12 ($24 + 12 + 12 + 12 = 60$), and is commercially available (Aldrich Chemical Company).

spheroids, and nitrobenzene molecules, each of which participates in face-to-face π - π interactions with two benzene moieties from two different calixarenes. Although guest species can be located within the interior of **1a** (that is, electron density maxima), it is not possible to determine their identity from the X-ray experiment, presumably owing to the high symmetry imposed by **1a** and high thermal motion within the cavity³.

Consultation of polyhedron models has revealed the structure of **1a** to conform to a snub cube (Fig. 2), one of the 13 Archimedean solids, in which the vertices of the square faces correspond to the corners of the calixarenes and the centroids of the eight triangles that adjoin three squares correspond to the eight water molecules¹⁰. Indeed, the ability of six calixarenes to self-assemble to form **1a** is reminiscent of spherical viruses in which identical copies of proteins self-assemble, by non-covalent forces, to form viral capsids with icosahedral cubic symmetry and a shell-like enclosure¹. In fact, **1a** is a closed-surface supramolecular spheroid that possesses $n > 3$ subunits and, owing to the fit exhibited by its components, possesses a topology that agrees with the theory of virus shell structure—which states that such octahedral systems must contain 24 asymmetric units and possess 4-3-2 symmetry¹. Moreover, such

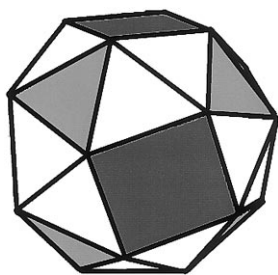


Figure 2 The snub cube, one of the 13 Archimedean solids. The square faces correspond to the calixarenes; the eight shaded triangles that adjoin three squares correspond to the water molecules of **1a**.

observations illustrate that to design related spherical hosts, one must consider the limited number of possibilities available in space for such frameworks, those being the five Platonic (regular) solids and the 13 Archimedean (semi-regular) solids¹⁰. It should be possible to select *a priori* appropriately sized, shaped and functionalized components to design spherical hosts in a way similar to the self-assembly exhibited by **1a**.

The hydrogen bond pattern that holds **1a** together is complex. Although we could not locate the hydroxy hydrogen atoms, it is possible to deduce the pattern knowing the positions and numbers of hydrogen-bond donors and acceptors. In principle, 64 hydrogen-bond donors and 112 hydrogen-bond acceptor sites are available to form **1a**. There are, however, two restrictions that arise from the positions of these sites that require the number of hydrogen bonds that define **1a** to equal 60 which, as a result, cause four hydroxy hydrogen atoms to point away from its surface. The first restriction lies in the positions of the water molecules. Their location on the three-fold axes limits the number of hydrogen bonds in which the water molecules can participate that contributes to **1a** to 24; in other words, each water molecule is capable of participating in only three hydrogen bonds along the surface of **1a**. The second restriction arises from the presence of the four intramolecular hydrogen bonds at the corners of the calixarenes. In particular, these interactions force four hydroxy groups to point their hydrogen atoms above each macrocycle, which as a consequence makes each calixarene a quadruple hydrogen bond donor¹¹. As each calixarene uses two hydroxy hydrogen atoms to form two hydrogen bonds to two calixarenes, this leaves two hydroxy hydrogen atoms per calixarene, which are subsequently used to form 12 hydrogen bonds to the water molecules. Consequently, to form the 12 hydrogen bonds, the water molecules, owing to their positions on the three-fold axes, must position four hydrogen atoms, each from a different water molecule, away from the surface of **1a**. To complete the hydrogen-bond array, the water molecules form 12 hydrogen bonds to 12 hydrogen-bond acceptor sites located on the calixarenes.

A closer inspection of the hydrogen bond pattern of **1a** reveals that each 'edge' of the water cuboid consists of a polar chain of five

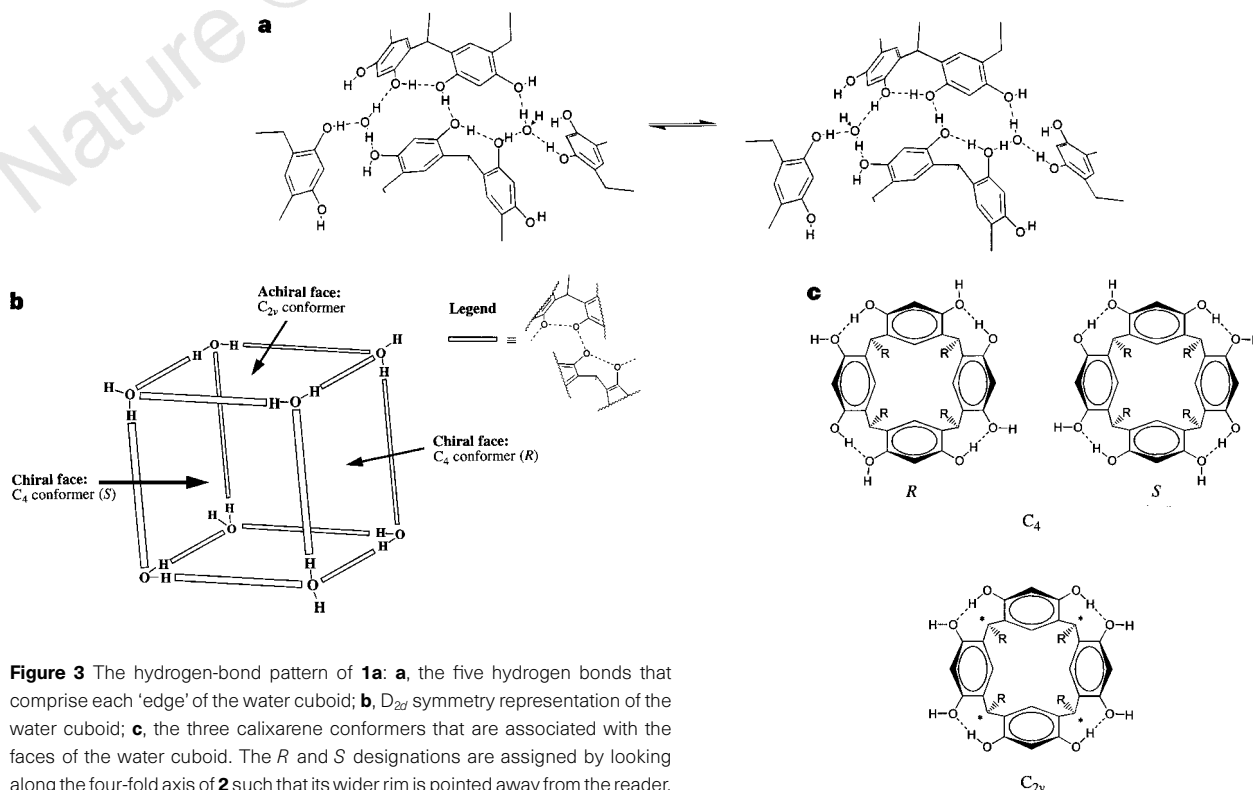


Figure 3 The hydrogen-bond pattern of **1a**: **a**, the five hydrogen bonds that comprise each 'edge' of the water cuboid; **b**, D_{2d} symmetry representation of the water cuboid; **c**, the three calixarene conformers that are associated with the faces of the water cuboid. The *R* and *S* designations are assigned by looking along the four-fold axis of **2** such that its wider rim is pointed away from the reader.

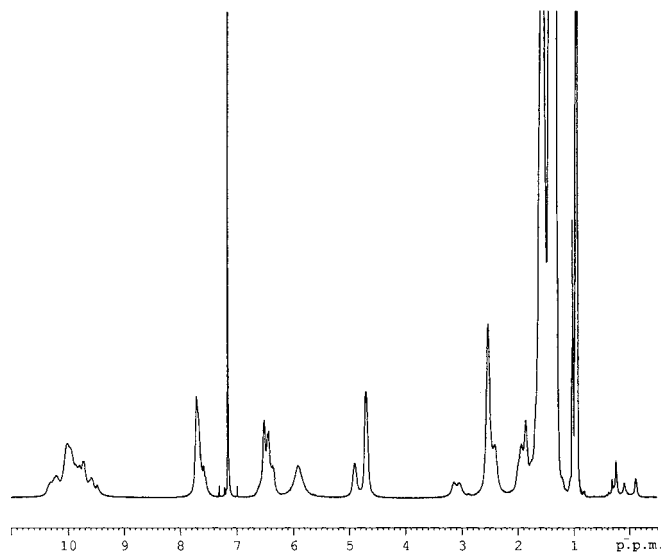


Figure 4 500 MHz ^1H NMR spectrum of **2b** in benzene- d_6 .

cooperative hydrogen bonds (Fig. 3a). As a result, each edge is capable of changing the sense of its direction by the interconversion of two water molecules from a hydrogen bond donor to a hydrogen bond acceptor and vice versa. Furthermore, as four water molecules point four hydrogen atoms away from the surface of **1a**, the edges of the cuboid are capable of undergoing such proton rearrangements simultaneously, which in effect makes **1a** a spherical proton pump able to channel the four 'dangling' protons to the eight corners of the water cuboid.

Although each of the dangling protons can be placed at any one of the eight corners of the water cuboid, it is most convenient, for symmetry reasons, to place them at the vertices of a tetrahedron such that the cuboid, ignoring the calixarenes, possesses D_{2d} symmetry (Fig. 3b)¹³. This arrangement furnishes the cuboid with four chiral faces and two achiral faces, which on completing the hydrogen bonds along each edge, gives rise to three distinct calixarenes, two of which are chiral (C_4) and one of which is achiral (C_{2v}) (Fig. 3c). Notably, the chiral calixarenes are associated with the chiral faces such that their handednesses match, whereas the achiral calixarenes are associated with the achiral faces. As a consequence of this arrangement, **1a** possesses D_2 symmetry.

Evidence supporting **2b** to maintain the structure of **1a** in solution has been obtained via one-dimensional and two-dimensional ^1H NMR measurements. Indeed, the spectrum of **2b** in benzene- d_6 , at increasing concentrations, displays resonances attributed to the chiral and achiral calixarenes as well as the eight water molecules (Fig. 4). In particular the CH hydrogens of the bridging methines, at 3.7 mM, exhibit two broad singlets at 4.69 and 4.90 p.p.m. (relative intensities 2:1) whereas their $\alpha\text{-CH}_2$ hydrogens exhibit a broad singlet at 2.52 p.p.m. and a pair of broad signals at 2.50 and 3.31 p.p.m. (relative intensities 4:1:1). These splittings and line-broadening effects are consistent with the presence of appreciably populated C_4 and C_{2v} conformers undergoing concerted hydrogen-bond tunneling motions in which the singlet at 4.90 p.p.m. and the signals at 2.50 and 3.31 p.p.m. are assigned to the C_{2v} conformer owing to the presence of chiral methine carbon atoms that induce desymmetrization of the methylene protons⁹. That the signals at 2.50 and 3.31 p.p.m. are coupled with the singlet at 4.90 p.p.m., as well as with each other, was verified by correlation spectroscopy NMR data. Furthermore, a broad singlet that shifts downfield with increasing concentration and integrates to 12 protons is observed at 5.91 p.p.m. and is assigned to the 12 protons of the water molecules that contribute to **1a**. This assignment is also supported by D_2O

exchange experiments. Indeed, these observations, coupled with a molecular mass determination in benzene ($7,066\text{ g mol}^{-1}$)¹⁴, provide convincing evidence supporting **2b** to maintain the structure of the spheroid in solution.

Our discovery that **2a** and **2b** form spherical molecular assemblies could bear relevance in a number of areas. In addition to providing insight into design criteria for the construction of spherical hosts, the fact that **1** sustains its cavity in both solution and the solid state suggests that it could be exploited to package guests within its interior^{5,9}. This could lead to a number of applications, which we are currently exploring, such as a chiral catalyst for chemical transformations⁸, a microvesicle for drug delivery¹⁵, and an intermediate for separations problems¹⁶. Indeed, molecular modelling experiments illustrate that the interior of **1** is spacious enough to accommodate relatively large substrates such as coordination compounds (for example ML_n $n = 1-6$), fullerenes (for example C_{60}), and supramolecular hosts (for example porphyrins). Considering the current commercial availability of **2a** and **2b**, **1** may now be recognized as a readily available spheroid in a burgeoning area of spherical host chemistry.

Methods

X-ray structure analysis. Crystal data for **3**: crystal size $0.10 \times 0.10 \times 0.10\text{ mm}$, cubic, space group $I432$, $a = 24.343(4)\text{ \AA}$, $U = 14425(4)\text{ \AA}^3$, $2\theta = 100^\circ$, Cu K α radiation ($\lambda = 1.54060\text{ \AA}$) for $Z = 2$. Least-squares refinement based on 691 reflections with $I_{\text{net}} > 2.0\sigma(I_{\text{net}})$ (out of 1,248 unique reflections) collected on an Enraf Nonius CAD-4 diffractometer equipped on a rotating anode generator and 148 parameters on convergence gave a final R of 0.124. Structure solution was accomplished with the aid of SHELXS-86 (ref. 17) and refinement was conducted using SHELXL93 (ref. 18) locally implemented on a Pentium-based IBM compatible computer. Many of the structures depicted in the figures were designed with the aid of RES2INS (ref. 19). □

Received 16 May; accepted 29 July 1997.

- Caspar, D. & Klug, A. Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.* **27**, 1-24 (1962).
- Kroto, H. W., Heath, J. R., O'Brien, S. C., Curl, R. F. & Smalley, R. E. C_{60} : Buckminsterfullerene. *Nature* **318**, 162-163 (1985).
- Casjens, S. in *Virus Structure and Assembly* (ed. Casjens, S.) 75-147 (Jones and Bartlett, Boston, 1995).
- Schwarz, H., Weiske, T., Böhme, D. K. & Hrusák, J. in *Buckminsterfullerenes* (eds Billups, W. E. & Ciufolini, M. A.) 257-283 (VCH, New York, 1993).
- Sherman, J. C. & Cram, D. J. Carcer and interiors provide a new phase of matter. *J. Am. Chem. Soc.* **111**, 4527-4528 (1989).
- Garel, L., Dutata, J.-P. & Collet, A. Complexation of methane and chlorofluorocarbons by cryptophane-A in organic solution. *Angew. Chem. Int. Edn Engl.* **32**, 1169-1171 (1993).
- Timmerman, P., Verboom, W., van Veggel, F. C., van Duynhoven, J. P. & Reinhoudt, D. N. A novel type of stereoisomerism in calix[4]arene-based carcerplexes. *Angew. Chem. Int. Edn Engl.* **33**, 2345-2348 (1994).
- Kang, J. & Rebek, J. Jr. Acceleration of a Diels-Alder reaction by a self-assembled molecular capsule. *Nature* **385**, 50-52 (1997).
- Shimizu, K. D. & Rebek, J. Jr. Synthesis and assembly of self-complementary calix[4]arenes. *Proc. Natl. Acad. Sci. USA* **92**, 12403-12407 (1995).
- Weninger, M. J. *Polyhedron Models* (Cambridge Univ. Press, New York, 1971).
- MacGillivray, L. R. & Atwood, J. L. Rational design of multi-component calix[4]arenes and control of their alignment in the solid state. *J. Am. Chem. Soc.* **119**, 6931-6932 (1997).
- Barbour, L. J. CAVITY, program for the determination of the volume of a molecular cavity. University of Missouri-Columbia, Missouri, USA (1997).
- Wales, D. J. & Ohmine, I. Rearrangements of model (H_2O)₈ and (H_2O)₂₀ clusters. *J. Chem. Phys.* **98**, 7257-7268 (1993).
- Aoyama, Y., Tanaka, Y. & Sugahara, S. Molecular recognition. 5. Molecular recognition of sugars via hydrogen-bonding interactions with a synthetic polyhydroxy macrocycle. *J. Am. Chem. Soc.* **111**, 5397-5404 (1989).
- Mathiowitz, E. et al. Biologically erodable microspheres as potential oral drug delivery systems. *Nature* **386**, 410-414 (1997).
- Atwood, J. L., Koutsantonis, G. A. & Raston, C. L. Purification of C_{60} and C_{70} by selective complexation with calixarenes. *Nature* **368**, 229-231 (1994).
- Sheldrick, G. M. Phase annealing in SHELX-90: direct methods for larger structures. *Acta Crystallogr. A* **46**, 467-473 (1990).
- Sheldrick, G. M. SHELXL93, program for crystal structure refinement, University of Göttingen, Germany (1993).
- Barbour, L. J. RES2INS, program for graphic preparation of SHELXL instruction files. University of Missouri-Columbia, Missouri, USA (1997).

Supplementary Information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from Mary Sheehan at the London editorial office of Nature.

Acknowledgements. We thank L. J. Barbour for writing the program CAVITY and K. T. Holman for synthesizing **2a**. This work was supported by the National Science Foundation (NSF) and the Natural Sciences and Engineering Research Council of Canada (NSERC) (research fellowship L.R.M.).

Correspondence and requests for materials should be addressed to J.L.A. (e-mail: chemja@showme.missouri.edu).

Correlation between rate of sea-level change and frequency of explosive volcanism in the Mediterranean

W. J. McGuire*, R. J. Howarth*, C. R. Firth†, A. R. Solow‡, A. D. Pullen§, S. J. Saunders†, I. S. Stewart† & C. Vita-Finzi*

* Greig Fester Centre for Hazard Research, Research School of Geological and Geophysical Sciences, University College London, Gower Street, London WC1E 6BT, UK

† Neotectonics Research Centre, Department of Geography & Earth Sciences, Brunel University, Isleworth, Middlesex TW7 5DU, UK

‡ Marine Policy Center, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

§ Department of Civil Engineering, Imperial College of Science, Technology and Medicine, Imperial College Road, London SW7 2BU, UK

Volcanic activity has frequently been linked to Quaternary environmental change, either by driving climate modification^{1,2} or in response to environmental changes³. Although a link has been established between large explosive eruptions and small (0.5 °C), brief (1–2 years) falls in global temperatures⁴, both the evidence and mechanisms responsible for longer episodes of eruption-induced planetary cooling remain questionable^{1,2,5,6}. In contrast, recent research based on ice-core data suggests that rapid climate changes during the past 110,000 years increased explosive volcanic activity⁷. Here we present a statistical analysis relating the frequency of explosive activity of Mediterranean volcanoes—based on dated^{8–11} tephra layers in deep-sea sediment cores—to the rate of late Quaternary sea-level change. The nonlinear correlation between the two is tentatively explained in terms of dynamic responses of the volcanoes to stress-related influences on various spatial scales. The correlation supports a mechanism or mechanisms by which the climate-driven growth and decay of large ice sheets can influence the eruptive chronologies of distant volcanic edifices via changes in global sea level.

The possibility that late Quaternary environmental changes influenced the frequency and magnitude of volcanic eruptions has

only recently been considered. New research based on the GISP2 ice core⁷ has produced a continuous record of explosive volcanism in the Northern Hemisphere over the past 110 kyr. The record identifies distinct periods of enhanced volcanic activity (notably at 35–22 and 17–6 kyr ago) which coincide with periods of rapid environmental change. In areas where active volcanism³ and Quaternary glaciation coincide, the correlation between the events can be explained by the effect of changing ice volumes on crustal stresses¹². In contrast, the effect of ice-sheet volume changes on unglaciated volcanic areas remains problematical. Several authors^{6,13–15} have proposed that meltwater loading and unloading could influence volcanic activity at sites distant from areas of ice accumulation through the global redistribution of water, although this hypothesis has never been tested. Here we examine the evidence for a link between such a redistribution and explosive volcanic activity, by comparing the temporal distribution of tephra layers in Mediterranean deep-sea cores^{8–11} with established global sea-level curves for the late Quaternary period. All the active volcanic centres in this region (Fig. 1) either form islands or are adjacent to coastlines, excepting Monte Vulture (Italy) which lies ~50 km from the sea.

The validity of both the correlation, and the proposed model, is crucially dependent on first, the accuracy with which ages are attributed to individual tephra layers, and second, on limited reworking and resedimentation of the tephra. The available data consist of 81 data layers (both megascopic and determined from a high percentage of volcanic glass in the sediment) in deep-sea cores from the Tyrrhenian^{10,11}, Adriatic¹¹, Ionian⁸ and Aegean^{8,9} seas, and the southeastern Mediterranean^{8–9} (Fig. 1). Paterne *et al.*¹⁶ argue convincingly in support of the completeness of the record in cores from the central Mediterranean. In this regard we are satisfied that sufficient accuracy (within a few kyr) is ensured by comparison of tephra layers with the oxygen-isotope record for each core, and by subsequent age calculation assuming constant sedimentation rates between two well-dated oxygen isotope signals in close proximity to one another. Additional temporal control is provided by correlation of a number of ash layers with well-known and dated terrestrial counterparts. This independent verification shows, importantly, that sedimentation rates are reasonably constant and unlikely to affect the validity of the tephra records. We accept, therefore, that variations in the numbers of tephra layers with time provide a true record of notable explosive activity at Mediterranean volcanoes during the late Quaternary period, given that the dominant wind directions were largely consistent over the period. The recent comparable correlation between tephra layers in the GISP2 ice core¹⁶ and periods of climate change during the late Quaternary provides additional support for the temporal distribution of the

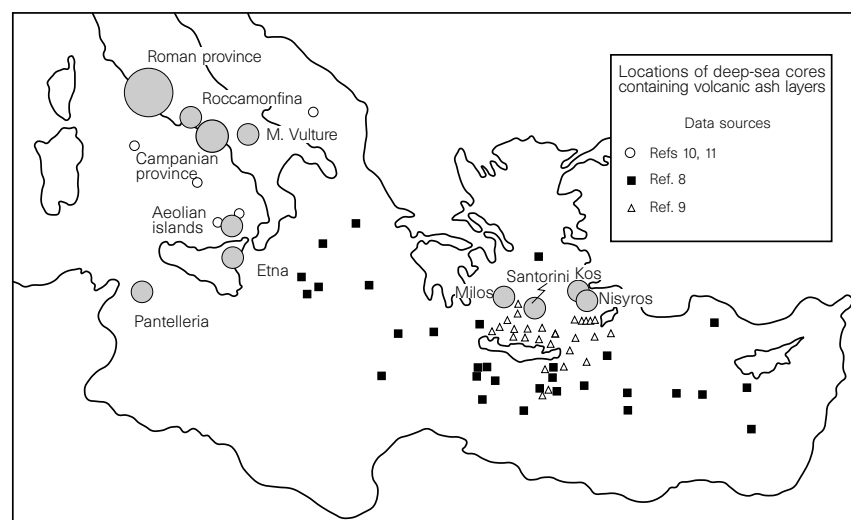


Figure 1 Location map of principal volcanic centres and provinces active in the Mediterranean region during the late Quaternary, and the distribution of boreholes from which deep-sea cores were extracted. The Roman Province includes the Vulcini, Vico, Sabatini and Albani centres; the Campanian Province includes Campi Flegrei, Somma-Vesuvius and Ischia.

tephra layers identified in the Mediterranean cores, and also, therefore, for limited reworking.

A cumulative plot of ordered event times, representing tephra-layer occurrence, versus time (Fig. 2) illustrates that the emplacement of detectable tephra layers in the submarine record has taken place, on average, every 1.05 kyr over the past 80 kyr, with the exception of three anomalous episodes of increased tephra-layer emplacement between 8 and 15, 34 and 38, and 55 and 61 kyr before present (BP), which have median repose periods (time to next tephra-producing event) of 0.35, 0.45 and 0.80 kyr, respectively. Paternite *et al.*^{17,18} reported an apparent periodicity of 22–24 kyr for the central Mediterranean data—tentatively attributed by them to glacio-eustatic influences on the plumbing system of the Campanian Volcanic Province—and our own spectral analysis supports their proposed periodicity. We propose, however, that the frequency of tephra-producing events and, by proxy, notable explosive eruptions at Mediterranean volcanoes, can be related to rapid variations in sea-level change during the late Quaternary period. In particular we draw attention to the quiescent phase (repose period 1.40 kyr) centred at 22 kyr ago and corresponding to the last low sea-level stand, and to the most intense episode in tephra-layer formation (repose period 0.35 kyr) between 15 and 8 kyr ago, which accompanied the very rapid rise in post-glacial sea levels¹⁹. Frequency modulation of the record of explosive eruptions at late Quaternary Mediterranean volcanoes by variations in the rate of sea-level change does not, however, require an increase in the overall magma output of the volcanoes involved, and as yet there is insufficient evidence to link rates of sea-level change with an elevated supply of magma from source.

If the underlying process governing the volcanic eruptions is a renewal process, that is, the intervals between eruptions are independent and identically distributed although not necessarily exponential (as in a Poisson process), then the goodness of fit of the observed sequence to a renewal model may be tested^{20,21}. If the process is either non-stationary or serially dependent, then the cumulative curve of Fig. 2 will wander away from the 45° line. The observed Kolmogorov–Smirnov and Cramér–von Mises statistics for departure of a scaled version of Fig. 2 from this null model were 0.123 and 0.002, respectively. These correspond to *P*-values (based on 10,000 randomizations of the observed interval times) of 0.23 and 0.47, respectively. The null hypothesis, that there is no evidence of a significant trend or serial dependence in the timing of the events, is therefore accepted. This supports the recent conclusion by Bebbington and Lal²² that use of non-homogenous models for volcanic eruptions is unwarranted.

We compare the sequence of inter-tephra repose periods with a smoothed curve of the rate of change of mean sea level (MSL) over the past 80 kyr. This was obtained by digitizing the data points in graphs of change in MSL with time from Barbados²³ and Pacific cores²⁴, which have been widely accepted as the best available MSL records. Figure 3a shows the estimated change in MSL as a function of age, together with the tephra layer record. The derived rate of change of MSL with time (based on 0.25-kyr intervals) is shown in Fig. 3b. The variation of the repose periods as a function of change of MSL with time is summarized in Fig. 4 (see legend for explanation). Comparison with 1,000 sets of simulated repose-time data with the same cumulative distribution as the observed data shows (Fig. 4) that the trend of the latter is statistically significant. This dependency has been corroborated by examining (1) the average rate of change of MSL, and (2) the average absolute rate of change of MSL, as a function of time following a tephra-producing event, using simply the slope of line segments connecting the original (that is, uninterpolated) observations in the MSL data. Comparison with empirical 90% tolerance intervals obtained using an interval-randomization procedure, confirms that in each case the rate of change of MSL tends to be negative in the vicinity of tephra-producing events. The critical dependence of the effect on MSL changes during

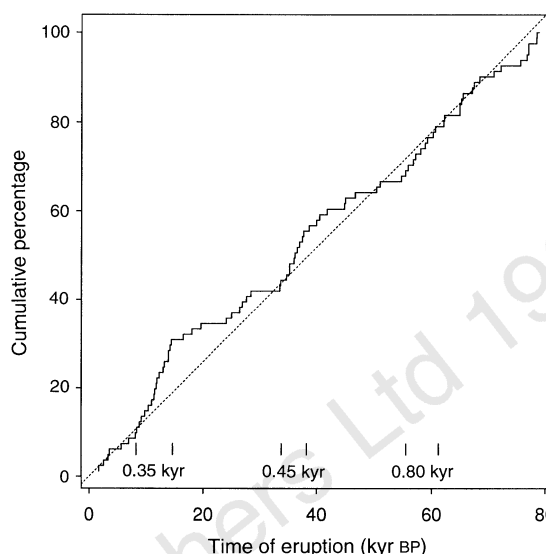


Figure 2 Cumulative plot of ordered event times (representing the tephra-layer occurrence) versus time. The dashed line corresponds to a median repose period of 1.05 kyr. Three anomalous episodes of increased tephra-layer emplacement between 8 and 15, 34 and 38, and 55 and 61 kyr BP are also shown, having median repose periods (time to next tephra-producing event) of 0.35, 0.45 and 0.80 kyr, respectively.

the past 15 kyr is unsurprising given the important sea-level changes involved in this post-glacial period.

The existence of a single causal link between the rate of sea-level change and the level of explosive activity is unlikely, and we suggest that a range of local or regional mechanisms are responsible, resulting from the dynamic responses of volcanoes to an array of stress-related influences. At the individual volcano scale, eruption may be triggered by a range of mechanisms including water-table changes, and variations in confining pressures. Finite-element analysis reveals, in fact, that large changes in sea level have contrasting effects on the internal stress regimes of coastal and island volcanoes. In the former case, loading due to a 100-m sea-level rise adjacent to the volcanic body reduces compressive stresses in the edifice by ~0.1 MPa, and may explain the triggering of explosive eruptions at 'charged' volcanoes in which bodies of relatively differentiated magma are stored at depths of 5 km or less. In contrast, expulsion of stored magma at island volcanoes is favoured by sea-level drawdown which reduces radial compressive stresses by as much as 1 MPa (for a 100-m fall). Furthermore, resulting shear-stress changes generated at the land–sea interface promote slope instability and the triggering of explosive eruptions by structural failure and collapse at volcanoes in both settings. The unique response of individual volcanoes to large changes in sea levels requires detailed study of each eruption record. This has already been accomplished for Etna, where the level of explosive eruptions is seen to fall to a marked low between 22 and 15 kyr ago, coincident with last low sea-level stand.

We speculate that broader-scale influences also occur, through slower-acting stress changes in continental margins and at island arcs. These may promote the ascent of fresh batches of magma into volcanoes, while increased levels of regional seismicity related to load redistribution may play a role in destabilizing already weakened volcanoes.

On a global scale the number of volcanoes susceptible to the above-mentioned effects is large. Current spatial distributions of active volcanoes show that 57% form islands or occupy coastal sites while a further 38% are located <250 km from a coastline. Assuming a similar distribution for around 1,500 volcanoes active during

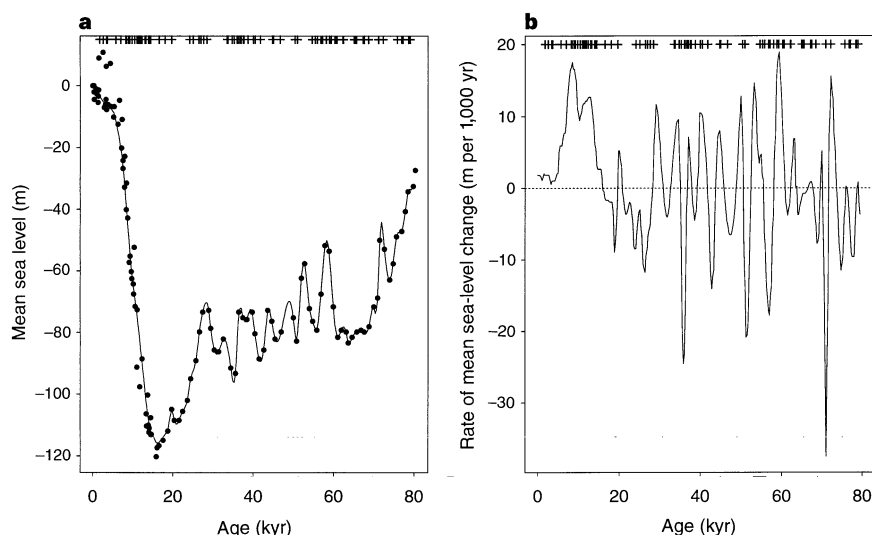


Figure 3 Changes in mean sea level. **a**, Estimated change in mean sea level (MSL) as a function of age (kyr) based on data from Barbados²³ and Pacific cores²⁴. A smooth curve has been fitted to the Barbados data and region of overlap of the two data sets using the non-parametric locally weighted regression smoother (LOWESS) technique²⁵. The sparser data of Shackleton²⁴ for the period 16–80 kyr ago have been fitted with a smooth cubic spline curve. Ages of dated^{8–11} tephra layers in deep-sea cores are shown by crosses. **b**, Rate of change of MSL with time, based on 25-kyr intervals. Ages of dated^{8–11} tephra layers in deep-sea cores are shown by crosses.

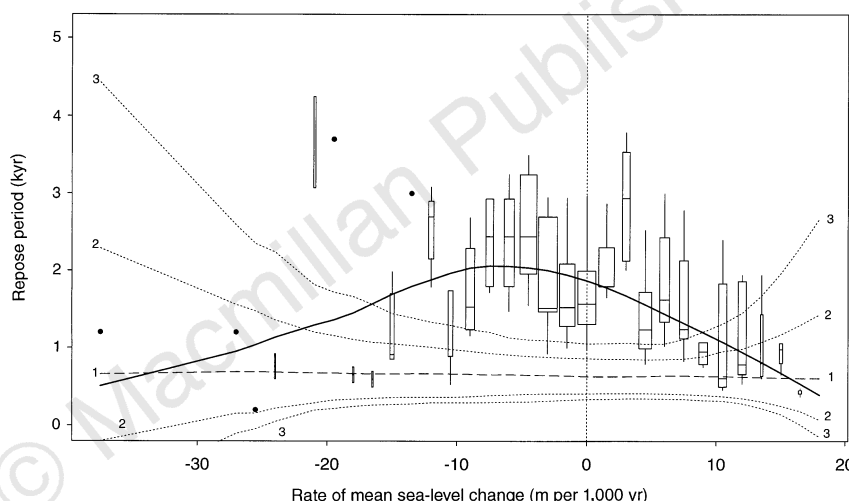


Figure 4 Variation of repose periods as a function of rate of change of mean sea level with time. These data are based on a bin width of $1.5d(\text{MSL})/dt$, and are summarized by box plots. Box width is proportional to the number of values in each bin; the base, horizontal dividing line, and top of each box show the 25th, 50th (median), and 75th percentiles. In a few cases the median coincides with the base or top of the box; whiskers extend out to the most extreme values lying within 1.5 of the interquartile range beyond the ends of the box. Isolated data points are shown individually. The bold, solid curve is a weighted LOWESS-smoothed fit²⁶ to the medians and indicates a clear decrease in repose period

with rate of change of MSL, either upwards (positive) or downwards (negative). We note that the maximum repose period is offset from the zero point on the rate of change axis, implying a time lag in the response of the volcanic systems to a given rate change in the sea-level record. The dashed lines show the median (line labelled '1') and empirical 95% ('2') and 99% ('3') confidence envelopes for the binning and LOWESS curve-fitting process applied to 1,000 sets of 81 repose times drawn randomly from the empirical cumulative distribution of the observed repose times. No systematic variation of repose period with rate of change of MSL is apparent in the simulated data.

the Holocene epoch, then $>1,400$ are likely to have been susceptible to the more direct effects of rapid sea-level change during the latter part of the Quaternary. Furthermore, the rapidity of these sea-level changes¹⁹, and consequently their potential to trigger responses in active volcanic structures, are only now becoming apparent. We thus expect similar correlations to hold true for many other volcanic provinces, and suggest that Quaternary environmental change, through changes in sea-level, has had a far bigger effect on the evolution of volcanic centres than has previously been suggested. □

Received 22 June 1995; accepted 11 August 1997.

1. Bray, J. R. Pleistocene volcanism and glacial initiation. *Science* **197**, 251–254 (1977).
2. Rampino, M. & Self, S. Volcanic Winter and accelerated glaciation following the Toba super-eruption. *Nature* **359**, 50–52 (1992).
3. Hall, K. Rapid deglaciation as an initiator of volcanic activity: an hypothesis. *Earth Surf. Process. Landforms* **7**, 45–51 (1982).
4. Dutton, E. G. & Christy, J. R. Solar radiative forcing at selected locations and evidence for global lower

- tropospheric cooling following the eruptions of El Chichón and Pinatubo. *Geophys. Res. Lett.* **19**, 2313–2316 (1992).
5. Ninkovitch, D. & Donn, W. L. Explosive Cenozoic volcanism and climatic implications. *Science* **194**, 899–906 (1976).
6. Rampino, M. R., Self, S. & Fairbridge, R. W. Can rapid climate change cause volcanic eruptions? *Science* **206**, 826–828 (1979).
7. Zielinski, G. A., Mayewski, P. A., Meeker, L. D., Whitlow, S. & Twickler, M. S. An 110,000-year record of explosive volcanism from the GISP2 (Greenland) ice core. *Quat. Res.* **43**, 109–118 (1996).
8. Keller, J., Ryan, W. B. F., Ninkovitch, D. & Altherr, R. Explosive volcanic activity in the Mediterranean over the past 200,000 yr as recorded in deep-sea sediments. *Geol. Soc. Am. Bull.* **89**, 591–604 (1978).
9. Vinci, A. Distribution and chemical composition of tephra layers from eastern Mediterranean abyssal sediments. *Mar. Geol.* **64**, 143–155 (1985).
10. Paterne, M., Guichard, F., Labeyrie, J., Gillot, P. Y. & Duplessy, J. C. Tyrrhenian sea tephrochronology of the oxygen isotope record for the past 60,000 years. *Mar. Geol.* **72**, 259–285 (1986).
11. Paterne, M., Guichard, F. & Labeyrie, J. Explosive activity of the south Italian volcanoes during the past 80,000 years as determined by marine tephrochronology. *J. Volcanol. Geotherm. Res.* **34**, 153–172 (1988).
12. Sigvaldason, G. E., Annertz, K. & Nilsson, M. Effect of glacier loading/deloading on volcanism: postglacial volcanic eruption rate of the Dyngjufjöll area, central Iceland. *Bull. Volcanol.* **54**, 385–392 (1992).
13. Nakada, M. & Yokose, H. Ice-age as a trigger of active Quaternary volcanism and tectonism. *Tectonophysics* **212**, 321–329 (1992).

14. Matthews, R. K. Tectonic implications of glacio-eustatic sea-level fluctuations. *Earth Planet. Sci. Lett.* **5**, 459–462 (1969).
15. Wallmann, P. C., Mahood, G. A. & Pollard, D. D. Mechanical models for correlation of ring-fracture eruptions at Pantelleria, Strait of Sicily, with glacial sea-level drawdown. *Bull. Volcanol.* **50**, 327–339 (1988).
16. Zielinski, G. A. *et al.* Volcanic aerosol records and tephrochronology of the Summit, Greenland, ice cores. *J. Geophys. Res.* (in the press).
17. Paterne, M., Labeyrie, J., Guichard, F., Mazaud, A. & Maitre, F. Fluctuations of the Campanian explosive volcanic activity (south Italy) during the past 190,000 years, as determined by marine tephrochronology. *Earth. Planet. Sci. Lett.* **98**, 166–174 (1990).
18. Paterne, M. & Guichard, F. Triggering of volcanic pulses in the Campanian area, south Italy, by periodic deep magma influx. *J. Geophys. Res.* **98**, 1861–1873 (1993).
19. Blanchon, P. & Shaw, J. Reef drowning during the last glaciation: evidence for catastrophic sea-level rise and ice-sheet collapse. *Geology* **23**, 4–8 (1995).
20. Cox, D. R. & Lewis, P. A. W. *The Statistical Analysis of Series of Events* 2nd edn (Methuen, London, 1978).
21. Solow, A. R. Testing for change in the frequency of El Niño events. *J. Clim.* **8**, 2563–2566 (1995).
22. Bebbington, M. S. & Lal, C. D. On homogeneous models for volcanic eruptions. *Math. Geol.* **28**, 585–600 (1996).
23. Fairbanks, R. C. A 17,000-year glacio-eustatic sea-level record: influence of glacial melting rates on the Younger Dryas event and deep-ocean circulation. *Nature* **342**, 637–642 (1989).
24. Shackleton, N. J. Oxygen isotopes, ice volume and sea level. *Quat. Sci. Rev.* **6**, 183–190 (1987).
25. Cleveland, W. S., Grosse, E. & Shyu, W. M. in *Statistical Models in S* (eds Chambers, J. M. & Hastie, T.) 309–376 (Wadsworth & Brooks/Cole, Pacific Grove, CA, 1992).

Acknowledgements. This work was supported by the Commission of the European Communities, DG XII, Environment Programme, Climatology and Natural Hazards Unit.

Correspondence and requests for materials should be addressed to W.J.McG. (e-mail: w.mcguire@ucl.ac.uk).

The effect of sedimentary cover on the flexural strength of continental lithosphere

Luc L. Lavier*† & Michael S. Steckler*

* Lamont-Doherty Earth Observatory, † Department of Earth and Environmental Sciences, Columbia University, PO Box 1000, RT 9W, Palisades, New York 10964–8000, USA

The factors that control the flexural rigidity—or effective elastic thickness (EET)—of continental lithosphere have been extensively studied over the past two decades. Using EET estimates derived from the analysis of topography, basin structures and gravity anomalies, several authors^{1–5} have shown that crustal thickness, geothermal gradient, strain rate, rheology and plate curvature all affect the flexural strength of continents. Recognition that certain combinations of these parameters result in a significant reduction of flexural strength caused by decoupling of the crust and the upper mantle^{3,5} has been a critical step in understanding why many continental areas have estimated EETs that are thin compared with the total mechanical thickness of the continental lithosphere⁵. Here we develop a semi-analytical model of the EET through a parametrization of the yield stress envelope^{6,7} that includes the effects of crust–mantle decoupling. We perform a detailed comparison of EET estimates at foreland basins and mountain belts to values predicted by our model and find that, to predict the EET estimates successfully, we need to take into account the effect of the sediment cover and to use a strong plagioclase-controlled rheology. The effect of sediment cover is to weaken the lithosphere because of the lower density of sediments relative to crystalline crust^{8,9} and by thermally insulating the lower crust^{9–11}.

The yield stress envelope (YSE)⁶ is one of the keys to understanding the variations of strength in the oceanic and continental lithosphere^{5,7,11}. It allows the parametrization of EET because it defines bounds on the elastic stresses that the lithosphere can sustain (Fig. 1). We formulate a model of flexural strength by assuming a double-layered brittle–elastic–plastic YSE for the continental lithosphere. In the brittle parts of the lithosphere, the yield stress follows Byerlee's law¹² (Fig. 1). By thin plate theory, the bending stress

within the elastic portion of the plate is linearly dependent on plate curvature (Fig. 1). The yield stress in the ductile lower crust and lower lithospheric mantle follows power-law dislocation creep^{12,13} represented by quasi-exponential curves (Fig. 1).

We derive semi-analytic expressions for the in-plane stress, T , and bending moment, M , for the flexing lithosphere by approximating the geotherm as a linear function of depth where the strength is controlled by ductile flow. The stress distribution, $\sigma(z)$, depends explicitly on the crustal thickness, the rheology of the continental lithosphere, the geotherm and the sediment thickness. Integrating, for a single layer we obtain:

$$T = \int \sigma(z) dz = \gamma \frac{1}{2} (z_1 - z')^2 + \dot{w}_0 \left[\frac{1}{2} (z_2^2 - z_1^2) - z_0 (z_2 - z_1) \right] + \left(\frac{\dot{\epsilon}}{A} \right)^{2/n} \frac{Q_A}{nRc} \int_{z_2}^{z_3} e^{u(z)} dz \quad (1)$$

where z_0 is the depth of the neutral surface; z_1 and z_2 represent the depth of the brittle–elastic boundary and the elastic–ductile boundary, respectively; and z_3 is the depth to the base of the YSE (Fig. 1). We have $z' = \{[\rho_c - \rho_s(z_s)]/\rho_c\}z_s$, where z_s is the depth of the bottom of the sediment layer and ρ_s and ρ_c are the average sediment and crustal densities, respectively. We have $\gamma = 0.67\bar{\rho}g$ or $\gamma = -2.21\bar{\rho}g$ for tension and compression, respectively, and $\dot{w}_0 = E/(1 - \nu^2)\dot{w}$, where $\dot{w}$ is the curvature of the lithosphere, E is Young's modulus and ν is Poisson's ratio. A is the power law pre-exponent, n the power law exponent and Q_A the activation energy. Finally, $u(z) = Q_A/[nR(T_0 + cz)]$, with T_0 and c as the zero intercept temperature and geothermal gradient for the best-fitting linear geotherm; R is the universal gas constant.

Equation (1) neglects some minor additional strength due to the variation of sediment density with depth. Similarly to the formulation in Bodine *et al.*⁷, the stress distribution, in-plane stress and bending moment are functions of the depths z_1 and z_2 . By specifying the curvature and in-plane stress, we solve equation (1) numerically to obtain z_1 and z_2 . Then we determine the stress distribution and calculate the bending moment, M :

$$M = \int \sigma(z) z dz \quad (2)$$

The flexural rigidity, D , can be calculated for any given value of the lithospheric curvature by using $M = -D\dot{w}$. Decoupling occurs

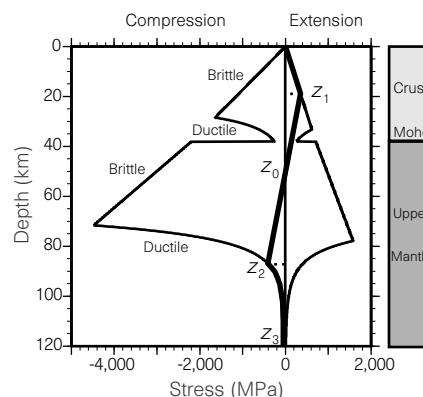


Figure 1 YSE of the continental lithosphere defining the maximum stress for which the lithosphere behaves like an elastic solid. The failure envelope of the crust and upper mantle forms a double layer, each consisting of a brittle upper part and ductile lower part. The heavy line shows an example of a bending stress profile. The lithosphere exhibits brittle failure from the surface to depth z_1 , linear elastic behaviour from z_1 to z_2 and ductile failure below depth z_2 ; z_0 corresponds to the depth of no elastic strain.

when bending stresses in the lower crust become larger than the yield stresses. Equation (1) is then modified and solved for a double YSE. The effective elastic thickness of the lithosphere is calculated with:

$$D = [E/12(1 - \nu^2)]EET^3 \quad (3)$$

For the modelling we use geotherms based on a 200 km equilibrium thickness for the lithosphere¹⁴. We calculate the geotherm for a given age of the lithosphere at the time of loading and a sediment thickness. We solve the heat equation by using a modified Fourier transform solution. The density and thermal conductivity of the sediments vary with depth according to compaction properties for a shaley sand¹⁵, a lithology representative of most foreland basins. The initial porosity is 56%, the compaction depth constant is 2,564 m and the thermal conductivity is $3.8 \text{ W m}^{-1} \text{ K}^{-1}$. Radiogenic heat generation is neglected for the sediments but is included for the crust. These values yield a conservative estimate of the effect of including the sediment cover. Loading by the sediment cover weakens the lithosphere in several ways: the low density of the sediments decreases the lithostatic pressure⁹; the low thermal conductivity of the sediment cover increases the temperature and decreases the yield stresses in the lower crust⁹; and the sedimentary layers increase the Moho depth, raising the lower crustal temperature.

In Fig. 2a we model the variation in EET caused by a change in crustal thickness for a continental lithosphere 1,000 Myr old at the time of loading. In agreement with Burov and Diamant⁵, a lithosphere with thick crust decouples for a smaller load or bending moment than its thermal equivalent with a thinner crust.

Figure 2b shows the effect of the sediment fill on flexural strength. For an average lithospheric structure similar to that of the Zagros foreland basin (Saudi Arabia–Iraq; see Table 1 and Fig. 3 for values), 5 km of sediment fill causes decoupling for curvatures $\geq 2 \times 10^{-7} \text{ m}^{-1}$. At this curvature the EET is reduced by 22 km. Comparison of the YSEs (Fig. 2b) highlights the effect of the sediment fill. In the absence of sediment fill, yield stresses in the lower crust are greater than the bending stresses and the lithosphere does not decouple. With sediments, yield stresses in the lower crust are smaller than the bending stresses and the lithosphere decouples, decreasing the EET.

We perform a detailed comparison of observed EET values estimated from studies of thrust belts at active and extinct plate boundaries to values predicted by our model (Table 1). To investigate the importance of weakening by sediment fill, we also consider a model parametrization that does not include the effects of the sediment cover (Fig. 3). The observed EET represents an average fit of the gravity anomaly across an area where the strength is variable. The average values are likely to be biased towards the dominant parts of the gravity anomaly signal. The lithospheric curvatures, estimated by fitting a circle to the depth of the base of the foreland basin sediments, are also an average that is biased towards the deeper part of the basin. Because they represent similar averages we can compare the predicted values of EET with the observed values of EET.

Our model successfully predicts the flexural rigidity of all of the examples except one (Fig. 3). We can discern three pools of values that illustrate the relative importance of sediments as a factor in controlling the flexural strength at continents. Where the litho-

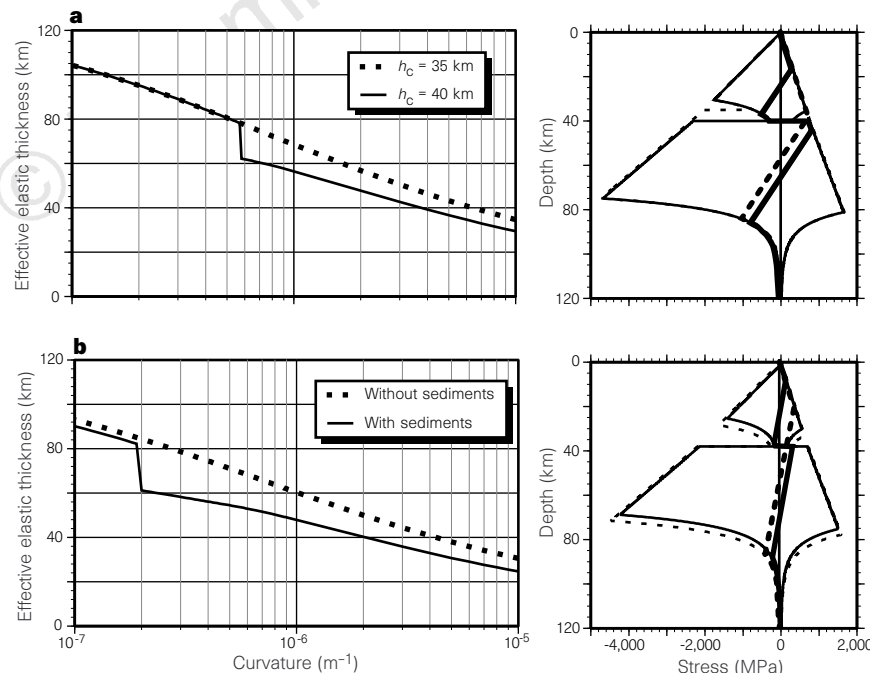


Figure 2 a, For 1,000-Myr-old lithosphere with a crust 35 km thick (dotted line), the EET decreases as curvature increases, but the lithosphere does not decouple. In the right panel, the corresponding YSE (broken light line) shows that the lower crust is always in the brittle field. Therefore, the lithosphere never decouples, as in the example bending stress profile with a curvature of $6 \times 10^{-7} \text{ m}^{-1}$. For an equivalent lithosphere with a 40-km crust (solid lines) the Moho depth is greater and thus the lower crust is hotter. The YSE (light solid line) for the lower crust exhibits a ductile failure zone. For curvatures greater than $6 \times 10^{-7} \text{ m}^{-1}$, the bending stresses (broken bold line) exceed the yield stress in the lower crust. This causes decoupling of the lithosphere. As a result, the bending stress profile readjusts (bold solid line) and the EET is reduced by $\sim 20 \text{ km}$. **b**, Plot of the

decrease in EET due to sediments. The 'with sediments' example (solid line) refers to a 600-Myr-old lithosphere with a 33-km crust and 5 km of sediments. 'Without sediments' (dotted line) refers to a similar lithosphere without sediments but with a 38-km crust to have an identical Moho depth. In the right panel, the YSE with a sediment cover is weaker than without a sediment cover owing primarily to the heating effect of the sediments. For the case without sediment cover, the EET decreases smoothly as the curvature increases. With sediment cover, the bending stresses in the lower crust exceed the yield stress of the lower crustal material for curvatures greater than $2 \times 10^{-7} \text{ m}^{-1}$; as shown by the stress profile (bold solid line). The lithosphere decouples and the EET value drops by $>20 \text{ km}$.

sphere is young and hot, and the crust sufficiently thick (Apennines, Eastern Alps, Western Alps), the lithosphere decouples regardless of the effect of the sediments. The parameters controlling the flexural strength are the age and the Moho depth. Where the lithosphere is old ($>1,000$ Myr) and cold (Ganges and Himalayas) and none of the weakening effects due to crustal thickness, curvature or sediment fill are sufficiently large, the lithosphere does not decouple. The flexural strength is controlled by the thickness and the rheology of the upper mantle. Finally, for the Tarim, Kunlun, Zagros, Southern Alps and Andes the age, the crustal thickness and the curvature alone are not enough to explain the observed EET values. There, the sediment fill plays a determining role by weakening an otherwise strong lithosphere. In these cases, the sediments cause a decrease of 8–30 km in the EET. The correct EET cannot be estimated without including the effect of the sediments.

Uncertainties on our modelling are described in Fig. 3. One source of uncertainty is critical to our modelling: the choice of rheological parameters for the lower crust^{12,13}.

For the lower crust we tested a set of weak rheologies (wet/dry quartzite and granite) and a set of stronger rheologies (diabase, anorthosite and quartz–diorite)¹³. The results show that when the creep is controlled by quartz rheology, the lithosphere is permanently decoupled⁵ even in very old lithosphere (Ganges and Himalayas). But plagioclase-controlled creep rheologies are sufficiently strong to maintain coupling at the Ganges and the Himalayas, as required by their high EET values. Quartz–diorite is the weakest possible rheology that allows such a behaviour. For stronger crustal rheologies, sediments are required to induce decoupling at other basins. Here we show the model with dislocation creep rheologies corresponding to anorthosite for the lower crust and olivine for the upper mantle. The constants, taken from Kirby and Kronenberg¹³, for anorthosite are $\log_{10}A = -3.5 \text{ MPa}^{-n} \text{ s}^{-1}$, $n = 3.2$, $Q_A = 238 \text{ kJ mol}^{-1}$, and for olivine are $\log_{10}A = 4.46 \text{ MPa}^{-n} \text{ s}^{-1}$, $n = 3.6$, $Q_A = 535 \text{ kJ mol}^{-1}$.

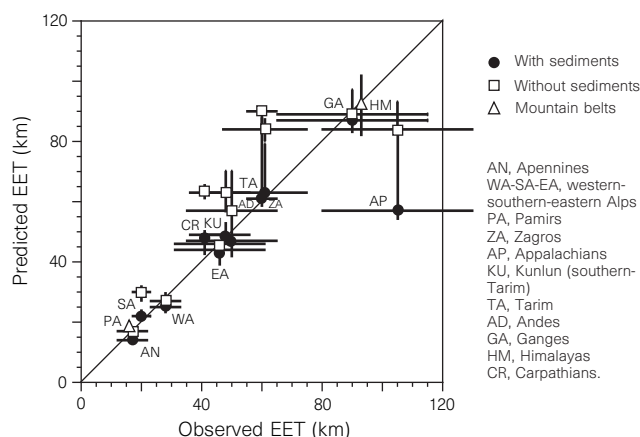
The Appalachian foreland is the one case in which we fail to predict the published EET of ~ 105 km (ref. 1). Several studies have

Table 1 Observations of EET and values used to calculate predicted EET

	D (10^{-23} N m)	Observed EET (km)	h_c (km)	$-\ddot{w}$ (10^{-7} m^{-1})	h_s (km)	Age (Myr)	Predicted EET (km)
Mountain belts							
Himalayas	~ 50	93 ± 15	60 ± 5	~ 3	–	$1,000 \pm 200$	90^{+10}_{-10}
Pamir	~ 0.1	16 ± 4	70 ± 5	~ 50	–	300 ± 100	18^{+2}_{-2}
Foreland basins							
Western Alps	1–3	28 ± 5	30 ± 3	40 ± 10	3–5	250 ± 50	25^{+2}_{-2}
Eastern Alps	4.6–6	46 ± 7	30 ± 3	10 ± 10	2–4	350 ± 100	44^{+2}_{-2}
Southern Alps	0.24–1.9	20 ± 3	25 ± 3	50 ± 10	5–7	250 ± 50	22^{+2}_{-2}
Apennines	~ 0.24	~ 17	24 ± 3	40 ± 10	3–5	100 ± 50	14^{+2}_{-2}
Carpathians	2–4	40.5 ± 5	32 ± 3	10 ± 10	5–7	$1,600 \pm 200$	49^{+2}_{-2}
Zagros	~ 10	~ 60	33 ± 3	2.5 ± 1	5–6	600 ± 200	61^{+2}_{-2}
Tarim	8.8–10	61 ± 14	35 ± 3	4 ± 1	5–7	800 ± 200	63^{+2}_{-2}
Kunlun	~ 5	~ 48	35 ± 3	15 ± 10	5–7	800 ± 200	49^{+2}_{-2}
Ganges	10–80	90 ± 25	35 ± 3	3 ± 1	3–5	$1,000 \pm 200$	87^{+2}_{-2}
Appalachians	10–100	105 ± 25	37 ± 3	2 ± 1	2–4	600 ± 200	60^{+2}_{-2}
Andes	2–11	50 ± 15	35 ± 3	20 ± 20	3–5	$1,000 \pm 200$	47^{+2}_{-2}

In this study we used observed EET values in areas for which we could obtain the input parameters (crustal thickness h_c , average curvature $\ddot{w}$, sediment thickness h_s , and thermal age at the time of loading) required for the model. The flexural rigidities were taken from ref. 5 and references therein and from refs 19–22. All values of observed EET were recalculated from the flexural rigidity, D , using $E = 5 \times 10^{10} \text{ N m}^{-2}$ and $\nu = 0.3$. The standard deviation of observed EET values were used to estimate uncertainties. We used only observed flexural rigidities derived from forward modelling of gravity anomalies because spectral method estimates remain controversial²³. The crustal thickness includes the thickness of sediments deposited before loading. The sediment thickness corresponds to the sediments deposited since the beginning of loading. The average curvature is derived from the horizon defining the base of h_s calculated on a spherical Earth. The age is the thermal age of the lithosphere at the time of loading. The uncertainties in h_c , h_s and $\ddot{w}$ correspond to uncertainties inherent in their measurement and to the variability observed in the cross-section from which the observed EET values were estimated. Large uncertainties in the average curvature are estimated owing to the limited data. We estimate the thermal age of the lithosphere as the difference between the age of the last major thermal event and the age of the beginning of collision. We assume uncertainties in the thermal age of 20–50% of the age determined from the literature. The references for crustal thickness, sediment thickness, the curvature and flexural rigidity are as follows: Zagros foredeep^{24,25}, Kunlun (South Tarim) foredeep^{26,27}, Ganges basin^{14,28}, Tarim foredeep^{28,30}, Andes foredeep^{29,30}, Appalachian foredeep^{31,32}, Carpathian foredeep^{22,23,33,34}, Apennine foredeep^{35,36}, Southern Alps^{4,37,38}, Eastern and Western Alps^{4,37,38} and Himalayas and Pamirs^{4,5,28,39}.

Figure 3 Plot of the predicted values of EET against observed values of EET. Perfect agreement is represented by the diagonal line. For the predicted values we used different values of strain rate. For active convergent areas (Himalayas, Ganges basin, Pamirs, Kunlun, Andes and Tarim) we used $\dot{\epsilon} = 10^{-15} \text{ s}^{-1}$. For fossil convergent margins (Alps, Appalachians, Apennines, Zagros and Carpathians) we used $\dot{\epsilon} = 10^{-17} \text{ s}^{-1}$. The white squares represent the predicted fit without taking into account the sediments deposited since the beginning of loading. The filled dots represent the predicted fit taking account of the effect of the sediments. The open triangles represent the predicted fit for the mountain belts (Pamirs and Himalayas). Error bars on the predicted values of EET are calculated as the means of the uncertainties due to uncertainties in the determination of h_s , h_c , $\ddot{w}$, the thermal age and the strain rate. Uncertainties in the predicted EET values due to uncertainties in the thermal age determinations vary between 2% and 10% for the areas with thermal ages greater than 600 Myr, and between 10% and 28% for areas with thermal ages from 100 to 350 Myr, the uncertainty being largest for the youngest age. The lithosphere is stronger at high strain rates than at low strain rates⁵ owing to the non-newtonian behaviour of power-law creep. An order of magnitude change in strain rate produces a change of 10–15% in the EET value. When the effect of the sediment cover is not included, an increase in curvature can cause decoupling for only two cases, TA and AP, as shown by the error bars for the predicted values. Overall, the model successfully predicts the EET values for most locations. The only observed value that is not fitted is the value for the Appalachian foredeep as discussed in the text.



not corroborated this high value^{16–18}. Quinlan and Beaumont¹⁶ matched the stratigraphy of the Appalachian basin by using a layered visco-elastic model that had an average EET of 67 km. Although they note that a pure elastic plate model does not adequately reproduce the stratigraphy, elastic plate models in which the EET decreases with curvature would produce offlapping stratigraphic patterns as seen in the observed stratigraphy. Recently, Stewart and Watts¹⁷ re-estimated the EET of several mountain belts by using a variable-rigidity formulation. Converting their estimates to the values of E and ν used in this paper gives a range of 50–88 km for the EET. These studies suggest that the high value found by Karner and Watts¹ is an overestimate and that a lower value averaging ~65–70 km is a better estimate. If so, our prediction of 60 km is within uncertainties.

We have established a parametrization of flexural strength at continents based on the yield stress envelope that successfully predicts the EET of the continental lithosphere at foreland basins and mountain belts. We have also demonstrated the importance of sediment fill as parameter controlling flexural strength at continents. The sediment cover is most likely to control the value of EET in places where the lithosphere's crust is thin compared with an average 35-km-thick continental crust⁵, the age of the lithosphere is close to its thermal equilibrium and for which the sediment cover reaches thicknesses greater than 3–5 km. Accounting for the effect of sediments and crustal thickness should facilitate the evaluation of the flexural strength at other types of basins and continental margins. □

Received 10 January; accepted 29 July 1997.

1. Karner, G. D. & Watts, A. B. Gravity anomalies and flexure of the lithosphere at mountain ranges. *J. Geophys. Res.* **88**, 10449–10477 (1983).
2. Kusznir, N. & Karner, G. D. Dependence of the flexural rigidity of the continental lithosphere on rheology and temperature. *Nature* **316**, 138–142 (1985).
3. McNutt, M. K., Diamant, M. & Kogan, M. G. Variations of elastic plate thickness at continental thrust belts. *J. Geophys. Res.* **93**, 8825–8838 (1988).
4. Royden, L. H. The tectonic expression slab pull at continental convergent boundaries. *Tectonics* **12**, 303–325 (1993).
5. Burov, E. G. & Diamant, M. The effective elastic thickness of the continental lithosphere: what does it really mean? *J. Geophys. Res.* **100**, 3905–3927 (1995).
6. Goetze, C. & Evans, B. Stress and temperature in the bending lithosphere as constrained by experimental rock mechanics. *Geophys. J. R. Astron. Soc.* **59**, 463–478 (1979).
7. Bodine, J. H., Steckler, M. S. & Watts, A. B. Observations of flexure and the rheology of the oceanic lithosphere. *J. Geophys. Res.* **86**, 3695–3707 (1981).
8. Cloetingh, S., Wortel, M. J. R. & Vlaar, N. J. Evolution of passive continental margins and initiation of subduction zones. *Nature* **297**, 139–142 (1982).
9. Steckler, M. S. & ten Brink, U. S. Lithospheric strength variations as a control on new plate boundaries: examples from the northern Red Sea region. *Earth Planet. Sci. Lett.* **79**, 120–132 (1986).
10. Karner, G. D. Sediment blanketing and the flexural strength of extended continental lithosphere. *Basin Res.* **3**, 177–185 (1991).
11. Molnar, P. & Tappin, P. A possible dependence of tectonic strength on the age of the crust in Asia. *Earth Planet. Sci. Lett.* **52**, 107–114 (1981).
12. Kohlstedt, D. L., Evans, B. & Mackwell, S. J. Strength of the lithosphere: constraints imposed by laboratory experiments. *J. Geophys. Res.* **100**, 17587–17602 (1995).
13. Kirby, S. H. & Kronenberg, A. K. Rheology of the lithosphere. *Rev. Geophys.* **25**, 1219–1244 (1987).
14. Zhang, Y.-S. & Tanimoto, T. High-resolution global upper mantle structure and plate tectonics. *J. Geophys. Res.* **98**, 9793–9823 (1993).
15. Sclater, J. G. & Christie, P. A. F. Continental stretching: an explanation of the post-Mid-Cretaceous subsidence of the central North Sea basin. *J. Geophys. Res.* **85**, 3711–3739 (1980).
16. Quinlan, G. M. & Beaumont, C. Appalachian thrusting, lithospheric flexure, and the Paleozoic stratigraphy of the Eastern interior of North America. *Can. J. Earth Sci.* **21**, 973–996 (1984).
17. Stewart, J. & Watts, A. B. Gravity anomalies and spatial variations of flexural rigidity at mountain ranges. *J. Geophys. Res.* **102**, 5327–5353 (1997).
18. Turcotte, D. L. & Schubert, G. *Geodynamics. Applications of Continuum Physics to Geological Problems*, 131–133 (Wiley, New York, 1982).
19. Snyder, D. B. & Barazangi, M. Deep crustal structure and flexure of the Arabian plate beneath the Zagros collisional mountain belt as inferred from gravity observations. *Tectonics* **5**, 361–373 (1986).
20. Watts, A. B., Lamb, S. H., Fairhead, J. D. & Dewey, J. F. Lithospheric flexure and bending of the central Andes. *Earth Planet. Sci. Lett.* **134**, 9–21 (1995).
21. Royden, L. H. & Karner, G. D. Flexure of lithosphere beneath Apennine and Carpathian foredeep basins: evidence for an insufficient topographic load. *Am. Ass. Petrol. Geol. Bull.* **68**, 704–712 (1984).
22. Macario, A., Malinverno, A. & Haxby, W. F. On the robustness of elastic thickness estimates obtained using the coherence method. *J. Geophys. Res.* **100**, 15163–15172 (1995).
23. Royden, L. H. & Burchfiel, B. C. Are systematic variations in thrust belt style related to plate boundary processes? (The western Alps versus the Carpathians). *Tectonics* **8**, 51–61 (1989).
24. Seber, D., Valve, M., Sandvol, E., Steer, D. & Barazangi, M. Middle East Tectonics: applications of Geographic Information Systems (GIS). *GSA Today* **7**(2), 1–6 (1997).
25. Koop, W. J. & Stoneley, R. Subsidence history of the Middle East Zagros Basin, Permian to Recent. *Phil. Trans. R. Soc. Lond. A* **305**, 149–168 (1982).
26. Fan, P. & Ma, B. L. *General Petroleum Geology of the Tarim Basin*, Vol. 1, 1–21 (Academia Sinica, Science Press, Beijing, 1990).
27. Teng, J. W. *Geophysical Fields and Hydrocarbon Prospects of the Tarim Basin*, Vol. 2, 24–40 (Academia Sinica, Science Press, Beijing, 1991).
28. Verma, R. K. Gravity field and nature of continent–continent collision along the Himalayas. *Phys. Chem. Earth* **18**, 385–403 (1991).

29. Wigger, J. W. et al. in *Tectonics of the Southern Central Andes, Structure and Evolution of an Active Continental Margin* (eds Reutter, K.-J., Scheuber, E. & Wigger, P. J.) 23–48 (Springer, Berlin, 1994).
30. Dunn, J. F., Hartshorn, K. G. & Hartshorn, P. W. in *Petroleum Basins of South America* (eds Tankard, A. J., Suárez Soruco, R. & Welsink, H. J.) 523–543 (Am. Ass. Petrol. Geol. Mem. 62, 1995).
31. Hinz, W. J. & Braille, L. W. in *The Geology of North America*, Vol. D-2 (Sedimentary Cover–North American Craton: U.S.) 5–24 (Geological Society of America, Boulder, CO, 1988).
32. Rankin, D. W. et al. in *Centennial Continent/Ocean Transect Vol. E-4 Central Kentucky to Carolina Trough 2 sheets* (Geological Society of America, Boulder, CO, 1991).
33. Roure, F., Roca, E. & Sassi, W. The Neogene evolution of the outer Carpathian flysch units (Poland, Ukraine and Romania): kinematics of a foreland/fold-and-thrust belt system. *Sediment. Geol.* **86**, 177–201 (1993).
34. Raileanu, V., Talos, D., Varodin, V. & Stipol, D. Crustal seismic reflection profiling in Romania on the Urziceni–Mizil line. *Tectonophysics* **223**, 401–409 (1993).
35. Bunn, H. in *Joint Interpretation of Geophysical and Geological Data Applied to Lithospheric Studies* (eds Giese, P. et al.) 193–215 (Kluwer, Dordrecht, 1991).
36. Ori, G. G. Continental depositional systems of the Quaternary of the Po plain (northern Italy). *Sediment. Geol.* **83**, 1–14 (1993).
37. Mueller, S. in *Mountain Building Processes* (ed. Hsue, K. J.) 181–199 (Academic, London, 1982).
38. Khale, H. G., Klinge, E., Mueller, S. & Egloff, R. The variation of crustal thickness across the Swiss Alps based on gravity and explosion seismic data. *Pure Appl. Geophys.* **114**, 479–494 (1976).
39. Belousov, V. V. et al. Structure of the lithosphere along the Deep Seismic Sounding profile: Tien Shan–Pamirs–Karakorum–Himalayas. *Tectonophysics* **70**, 193–221 (1980).

Acknowledgements. We thank B. Coakley, C. Ebinger, R. Newman and N. White for constructive comments that improved the paper. We thank the NSF Earth Sciences, and F. Brigaud and Elf-Aquitaine Exploration Production, for support.

Correspondence and requests for materials should be addressed to L.L.L. (e-mail: luc@ldeo.columbia.edu).

Failure of plume theory to explain midplate volcanism in the southern Austral islands

M. K. McNutt*, D. W. Caress†, J. Reynolds‡, K. A. Jordahl*§ & R. A. Duncan||

* Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

† SeaBeam Instruments, 141 Washington Street, East Walpole, Massachusetts 02032, USA, and Lamont-Doherty Earth Observatory, Palisades, New York 10964, USA

‡ Monterey Bay Aquarium Research Institute, PO Box 628, Moss Landing, California 95039, USA

§ MIT/Woods Hole Oceanographic Institute Joint Program in Oceanography and Applied Ocean Science and Engineering, Woods Hole, Massachusetts 02543, USA

|| College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331, USA

It has long been recognized that the properties of the Cook–Austral chain (Fig. 1) of volcanoes in the South Pacific are difficult to reconcile with the theory that volcanic activity in plate interiors is produced by the drift of tectonic plates over narrow, stationary plumes¹ of hot mantle material upwelling from depth. Radiometric dates^{2,3} from many island samples are younger or older than would be predicted if a single plume currently located at volcanically active Macdonald seamount⁴ was responsible for all of the volcanoes. Indeed, only the southernmost part of the Austral volcanic line has hitherto appeared to be consistent with plume activity, and then only within the past 6 million years (Myr)^{5,6}. Here we report radiometric dates that demonstrate that these southern Austral volcanoes are actually composed of three distinct volcanic chains with a range of ages spanning 34 Myr and with inconsistent age progressions. Gravity anomalies and sea-floor fabric suggest that the volume and location of volcanism in this region is controlled by stress in the lithosphere rather than the locus of narrow plumes rising from the deep Earth.

Our data were collected during a two-month expedition to the southeastern end of the Austral chain of volcanoes (Fig. 2) aboard the RV *Maurice Ewing* in March–May 1996. Our detailed mapping reveals that, in addition to the well-known Macdonald chain, there are two other lines of volcanoes, which we have named the

Ngatemato and Taukina chains after the ruling families of Rapa. Radiometric dating (^{40}Ar – ^{39}Ar incremental heating methods) of rocks dredged from these volcanoes (Table 1) indicates that the Ngatemato chain erupted more than 30 Myr ago on lithosphere that was only 10 Myr old at the time. The volcanoes consist of low, *en échelon* ridges of slightly enriched tholeiitic basalts which differ dramatically from the radial rift zones and extremely alkaline rocks on volcanoes of the Macdonald chain. The old age of the Ngatemato chain was further confirmed by the low reflectance of the sonar signal returned to the Hydrosweep mapping system, indicating considerable burial beneath a blanket of pelagic sediment. The small Taukina chain erupted only slightly later, and is composed of small, tholeiitic, ‘haystack-shaped’ volcanoes frequently displaying summit calderas. It is surprising that both young (0–5 Myr) and

old (29–32 Myr) dates are found for rocks dredged from the deep flanks of the Macdonald line, indicating at least two distinct periods of edifice formation. When our dated samples are reconstructed to their positions at the time of volcanism, using rotation poles⁷ determined for motion of the Pacific plate with respect to fixed mantle plumes (Fig. 1), we find that there would have to be nearly as many ‘plumes’ as there are dated samples. The reconstructions pass through the Foundation seamounts⁸, indicating a long-lived but sporadic near-ridge enhancement in melting.

Gravity data can be used to estimate the approximate volumes of young and old volcanism in the southern Austral islands by using the established correlation between the elastic strength of the oceanic lithosphere, as measured by its effective elastic thickness T_e , and the age of the oceanic lithosphere at the time of volcanism⁹.

Table 1 ^{40}Ar – ^{39}Ar age determinations for whole rock basalts

Sample	Name	Chain	Latitude	Longitude	Plateau		Total fusion age (Myr)
					^{39}Ar (%)	Age (Myr)	
EW96 18-1	Make	N	28° 32' S	140° 13' W	79	25.58 ± 1.01	27.7
EW96 13-2	Aureka	N	28° 12' S	141° 13' W	100	31.30 ± 0.74	31.3
EW96 13-1	Aureka	N	28° 12' S	141° 13' W	No plateau developed		28.8
EW96 19-1	Evelyn	T	27° 41' S	139° 25' W	100	25.95 ± 1.15	26.1
EW96 19-4	Evelyn	T	27° 41' S	139° 25' W	No plateau developed		24.2
EW96 20-6	Herema	T	27° 28' S	140° 00' W	100	22.47 ± 1.48	23.4
EW96 9-13	Opu	T	27° 02' S	143° 09' W	65	33.94 ± 0.62	39.1
EW96 9-25A	Opu	T	27° 02' S	143° 09' W	No plateau developed		31.7
EW96 14-1	Ra	M	28° 46' S	141° 07' W	88	29.21 ± 0.61	27.5
EW96 7-5	Marotiri	M	27° 57' S	143° 36' W	54	31.95 ± 0.82	39.6
EW96 7-7	Marotiri	M	27° 57' S	143° 36' W	100	3.78 ± 0.18	3.75

Ages are reported relative to biotite monitor FCT-3 (28.03 ± 0.18 Myr), which is calibrated against hornblende Mmhb-1 (523.5 Myr). Plateau ages are the mean of concordant step ages (3 or 4 steps for each sample that developed a plateau), weighted by the inverse of their variances. Calculations use the following decay and reactor interference constants: $\lambda_{\text{e}} = 0.581 \times 10^{-10} \text{ yr}^{-1}$; $\lambda_{\text{g}} = 4.963 \times 10^{-10} \text{ yr}^{-1}$; $(^{36}\text{Ar}/^{39}\text{Ar})_{\text{Ca}} = 0.000264$; $(^{38}\text{Ar}/^{39}\text{Ar})_{\text{Ca}} = 0.000673$; $(^{40}\text{Ar}/^{39}\text{Ar})_{\text{Ca}} = 0.01$. The volcanic chain to which each seamount belongs is identified by M, N or T for the Macdonald, Ngatemato and Taukina chains, respectively.

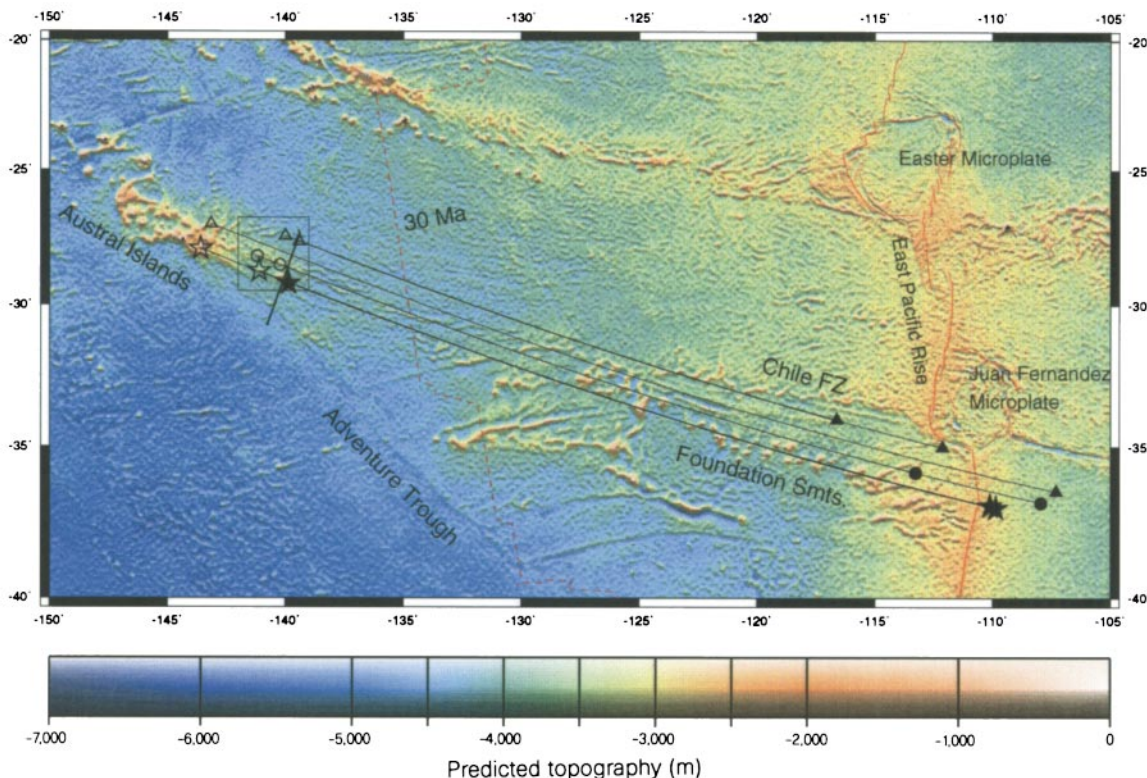


Figure 1 Bathymetric map of the South Pacific seen from ship soundings and satellite altimetry²⁴. The solid red line is the locus of present-day sea-floor spreading in the Pacific. The broken red line is the 30-Myr isochron²⁵. The box shows the location of the region surveyed in the southern Austral islands. Open circles (Ngatemato), triangles (Taukina) and stars (Macdonald) indicate the

location of dredge sites for dated volcanic rocks; filled counterparts show the reconstructed positions of the dated volcanic rocks when they erupted using Pacific-hotspot finite rotation poles⁷. The thick black line gives the location of the profile shown in Fig. 3.

We begin at 30 Myr by Ngatemato volcanic load 4.5 km high erupting, approximately as a gaussian function (Fig. 3a), on a young, hot, weak plate ($T_e = 5$ km). This load causes a moat 3.3 km deep to form, such that the net height of the volcano is only 1.2 km (Fig. 3b). The Macdonald volcanism then erupts a small, volcanic load 1 km high 30 Myr later (Fig. 3c) on an older, stiffer ($T_e = 15$ km) plate. The moat it creates is 630 m deep

(Fig. 3d), but because the volcano rests on the flanks of the larger, older volcano, the net height after plate bending is still 800 m. The older Ngatemato volcano is also affected by this younger volcanism, and becomes even further submerged in the moat created by the younger load, thus reducing its apparent height by another few hundred metres. The net effect of this simple modelling exercise is a bimodal seamount chain (Fig. 3d) with the Macdonald volcano

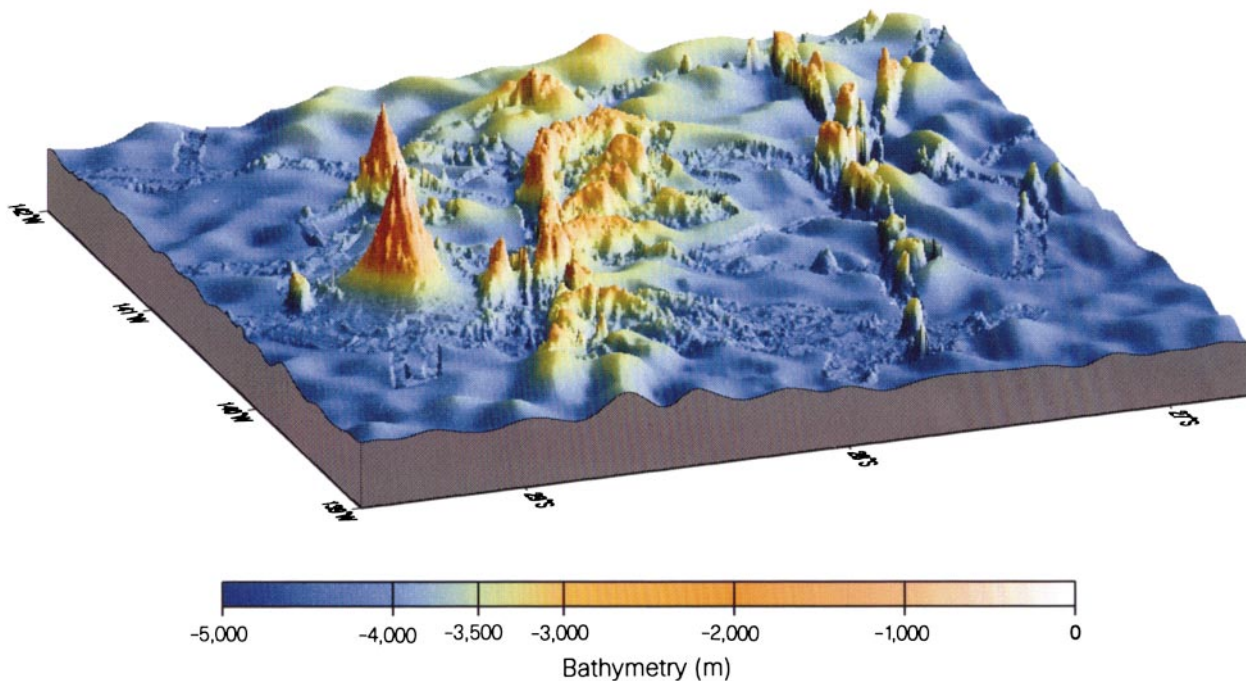


Figure 2 Perspective view of the southern Austral islands showing the three distinct lines of volcanoes. High-resolution swath bathymetry has been overlaid on a base map of bathymetry predicted from satellite altimetry²⁴, as there were too few ship soundings in this area before our expedition to determine even the grossest contours. The view is from the southeast, with illumination from the

north. The Ngatemato seamounts (centre) are low, *en échelon* ridges with summit calderas, whereas the Taukina seamounts (right) are individual cones; both have a convex shape. In contrast, the Macdonald seamounts (left) display prominent rifts and more concave slopes.

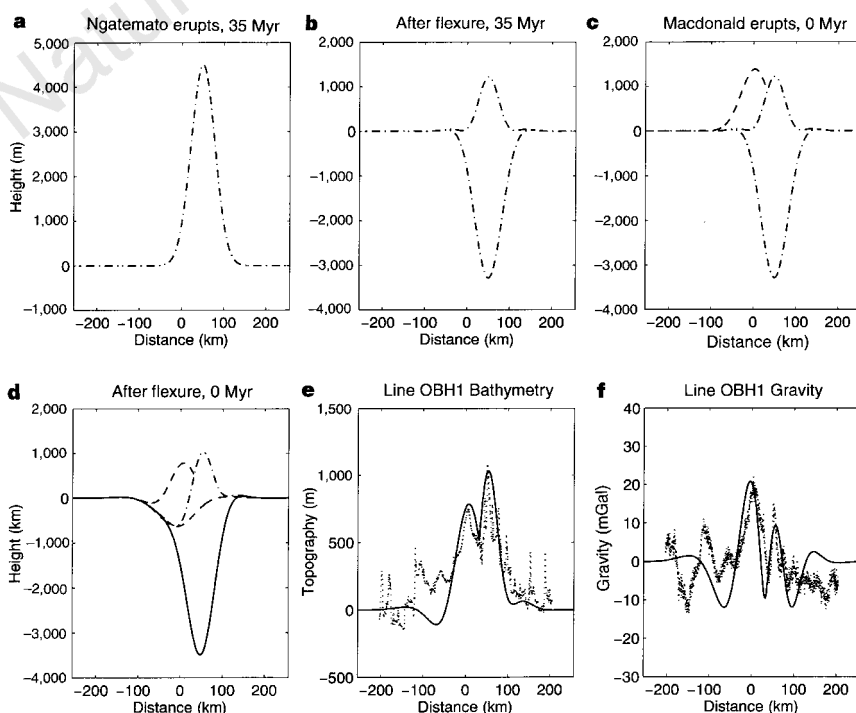


Figure 3 Flexural modelling of lithospheric loading at two distinct times. **a**, Load from Ngatemato volcanism. **b**, Ngatemato load subsided after lithospheric flexure. **c**, Load from Macdonald volcanism (dashed line) added to flank of Ngatemato load. **d**, After lithospheric flexure from Macdonald load (dashed line). Solid line is the combined moat from both loads. **e**, Net predicted topography (solid line) compared to observed topography (dots) along line OBH1 in Fig. 1. **f**, Net predicted gravity anomaly (solid line) obtained from the density effects of the topography in (solid line in **e**) and the depression of the crust-mantle boundary predicted from the plate bending in (solid line in **d**) compared with the observed gravity anomaly (dots).

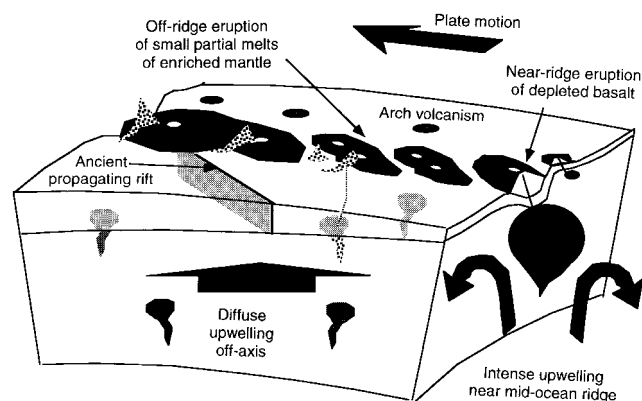


Figure 4 Schematic representation of a model for volcanism in the Austral islands. We assume that the upper mantle beneath French Polynesia is enriched in several different easily melted components that have been reinjected into the upper mantle by earlier subduction of oceanic crust, continental crust, and sediments. Their unusual isotopic signature would be homogenized and diluted by the large degrees of partial melting that occur near the mid-ocean ridge but preserved in the smaller off-ridge seamounts. These more exotic melts are pervasively available beneath the plate on account of diffuse upwelling, but are preferentially channelled to the surface by pre-existing cracks in the plate, topography on the base of the plate, and older volcanic conduits.

smaller than the Ngatemato volcano, and a net gravity anomaly (the difference between the mass excess of the volcanoes and the mass deficit caused by their combined flexural signals) that is larger over the Macdonald chain (Fig. 3e), owing to the asymmetrical crustal flexure across the section.

Other profiles measured confirm that most of the Macdonald chain is young (0–5 Myr) and most of the Ngatemato chain is old (~30 Myr). The effective elastic thickness of the plate supporting Marotiri to the northwest is only 10 km, as opposed to the expected 15 km if the entire feature were only 5 Myr old, suggesting that there is more mixing of volcanic ages in the northern section of the Macdonald chain of volcanoes. We believe that this complicated interaction of the flexural effects of volcanoes of different ages is responsible for the anomalously low effective elastic plate thicknesses reported for the South Pacific^{10–12}. Young volcanoes are not loading an old, anomalously weak plate; rather, the majority of the load erupted when the plate was young.

The flexure models suggest a causal mechanism for the spatial association of these three chains spanning a 34 Myr age range. The Taukina chain lies along the flexural arch of the plate bending produced by the Ngatemato chain. Arch volcanism is well documented in Hawaii, and presumably indicates that either the extensional stress in the upper elastic plate facilitates and focuses magma outpouring, and/or that the arch provides a structural trap¹³ for low-density magma (Fig. 4). The Macdonald chain also erupts where the lithosphere is in tension from Ngatemato loading, although somewhat inboard of the arch maximum. The volume of all three chains is also modulated along strike (Fig. 4), with maximum volumes corresponding to the intersection of the north-northwestern line with the Adventure trough, an ancient propagating rift¹⁴.

Other morphological evidence from the southern Austral islands also suggests a wide range of ages for volcanism, which is inconsistent with plume theory. Deeply submerged, flat-topped seamounts¹⁵ are interspersed with young volcanoes, suggesting erosion at wave base during a much earlier phase of volcanic activity when the sea floor was shallower. Furthermore, individual islands within the long chain display totally different isotopic signatures, thought to be characteristic of different mantle source regions¹⁶, making the hypothesis of a single plume even more unlikely.

Modifications to plume theory that might explain the Cook–Austral chain include proposals such as fortuitous alignment of several distinct plumes¹⁷ or volcanic rejuvenation that repaves older features with a veneer of young rock long after it has passed over the plume¹⁸. Our data suggest that such minor adjustments to plume theory are inadequate.

The picture that emerges instead has the eruption of the Ngatemato chain as the primary event, beginning more than 30 Myr ago by excess melt production near a mid-ocean ridge. The locus of later volcanism was controlled by stresses in the lithosphere caused by Ngatemato loading and modulated by another ancient plate structure, the Adventure trough. Seismic tomographic imaging of the upper mantle of the South Pacific¹⁹ and dynamic modelling²⁰ demonstrate that the upper mantle beneath this entire region is anomalously hot and weak, suggesting there were unusual amounts of partial melt and broad-scale upwelling. The geochemistry of the mantle melts is more consistent with sampling an upper mantle enriched with easily melted components through plate subduction than with transport of primitive mantle from deeper regions that never participated in the plate tectonic cycle^{21–23}. The plate structures merely organize the outflow of this easily melted component entrained in the diffuse upwelling into the chains we observe today. Shearing between the drifting plate and the underlying melt source would lead to extremely short age progressions of volcanoes tapping an individual melt pocket. Although we cannot definitively rule out the existence of a plume beneath the Macdonald seamount, if one exists in this region it is responsible for only a small amount of the very substantial volume of off-ridge volcanism. □

Received 24 February; accepted 7 July 1997.

1. Morgan, W. J. Convection plumes in the lower mantle. *Nature* **230**, 42–43 (1971).
2. Dalrymple, G. B., Jarrard, R. D. & Clague, D. A. K–Ar ages of some volcanic rocks from the Cook and Austral Islands. *Geol. Soc. Am. Bull.* **86**, 1463–1467 (1975).
3. Turner, D. L. & Jarrard, R. D. K–Ar dating of the Cook–Austral Island chain: A test of the hot spot hypothesis. *J. Volcanol. Geotherm. Res.* **12**, 187–220 (1982).
4. Johnson, R. H. Active submarine volcanism in the Austral Islands. *Science* **167**, 977–979 (1970).
5. Krummenacher, D. & Noetzelin, J. Ages isotopiques K/Ar de roches prélevées dans les possessions françaises du Pacifique. *Soc. Géol. Fr. Bull.* **8**, 173–175 (1966).
6. Duncan, R. A. & MacDougall, I. Linear volcanism in French Polynesia. *J. Volcanol. Geotherm. Res.* **1**, 197–227 (1976).
7. Duncan, R. A. & Clague, D. A. in *The Pacific Oceans* (eds Nairn, A. E. M., Stehli, F. G. & Uyeda, S.) (Plenum, New York, 1985).
8. Mammerickx, J. The Foundation Seamounts: tectonic setting of a newly discovered seamount chain in the South Pacific. *Earth Planet. Sci. Lett.* **113**, 293–306 (1992).
9. Watts, A. B., Bodine, J. H. & Ribe, N. R. Observations of flexure and the geological evolution of the Pacific basin. *Nature* **283**, 532–537 (1980).
10. McNutt, M. K. & Menard, H. W. Lithospheric flexure and uplifted atolls. *J. Geophys. Res.* **83**, 1206–1212 (1978).
11. Calmant, S. & Cazenave, A. Anomalous elastic thickness of the oceanic lithosphere in the south-central Pacific. *Nature* **328**, 236–238 (1987).
12. Goodwillie, A. M. & Watts, A. B. An altimetric and bathymetric study of elastic thickness of the central Pacific Ocean. *Earth Planet. Sci. Lett.* **118**, 311–326 (1993).
13. Sleep, N. H. Lateral flow and ponding of starting plume material. *J. Geophys. Res.* **102**, 10001–10012 (1997).
14. Cande, S. C. & Haxby, W. F. Eocene propagating rifts in the southwest Pacific and their conjugate features on the Nazca plate. *J. Geophys. Res.* **96**, 19609–19622 (1991).
15. Johnson, R. H. in *Research Reports* (eds Oehser, P. H., Lea, J. S. & Powars, N. L.) 389–405 (National Geographic Society, Washington DC, 1971).
16. Hart, S. R. A large-scale isotope anomaly in the southern hemisphere mantle. *Nature* **309**, 753–757 (1984).
17. Barszczus, H. G. *Les îles Australes (Polynésie Française) et la théorie des points chauds* (Office de la Recherche Scientifique et Technique Outre-Mer, 1980).
18. Menard, H. W. & McNutt, M. K. Evidence for and consequences of thermal rejuvenation of the lithosphere. *J. Geophys. Res.* **87**, 8570–8580 (1982).
19. Su, W.-J. *The Three-Dimensional Shear-Wave Velocity Structure of the Earth's Mantle* (Harvard Univ. Press, 1992).
20. McNutt, M. K. & Judge, A. V. The superswell and mantle dynamics beneath the South Pacific. *Science* **248**, 969–975 (1990).
21. Hedge, C. E. Strontium isotopes in basalts from the Pacific Ocean basin. *Earth Planet. Sci. Lett.* **38**, 88–94 (1978).
22. Richter, F. M. & Ribe, N. M. On the importance of advection in determining the local isotopic composition of the mantle. *Earth Planet. Sci. Lett.* **43**, 212–222 (1979).
23. Hofmann, A. W. Mantle geochemistry: the message from oceanic volcanism. *Nature* **385**, 219–229 (1997).
24. Smith, W. H. F. & Sandwell, D. T. New global seafloor topography from satellite altimetry. *Eos* **77**, F315 (1996).
25. Müller, R. D., Roest, W. R., Royer, J.-Y., Gahagan, L. M. & Sclater, J. G. Digital isochrons of the world's ocean floor. *J. Geophys. Res.* **102**, 3211–3214 (1997).

Acknowledgements. We thank W. Smith for providing the predicted bathymetry data, and J.-G. Schilling, M. Maia and N. Sleep for reviews. This work was supported by the US National Science Foundation.

Correspondence should be addressed to M.K.M. (e-mail: mcnut@mit.edu).

Epipubic bones in eutherian mammals from the Late Cretaceous of Mongolia

Michael J. Novacek*, Guillermo W. Rougier*, John R. Wible†, Malcolm C. McKenna*, Demberelyin Dashzeveg‡ & Inés Horowitz*

* Department of Vertebrate Paleontology, American Museum of Natural History, New York, New York 10024, USA

† Department of Anatomical Sciences and Neurobiology, School of Medicine, University of Louisville, Louisville, Kentucky 40292, USA

‡ Geological Institute, Mongolian Academy of Sciences, Ulaan Baatar, Mongolia

An important transformation in the evolution of mammals was the loss of the epipubic bones. These are elements projecting anteriorly from the pelvic girdle into the abdominal region in a variety of Mesozoic mammals, related tritylodonts, marsupials and monotremes but not in living eutherian (placental) mammals^{1–3}. Here we describe a new eutherian from the Late Cretaceous period of Mongolia, and report the first record of epipubic bones in two distinct eutherian lineages. The presence of epipubic bones and other primitive features suggests that these groups occupy a basal position in the Eutheria. It has been argued that the epipubic bones support the pouch in living mammals^{1,3,4}, but epipubic bones have since been related to locomotion and suspension of the litter mass of several attached, lactating offspring⁵. The loss of the epipubic bones in eutherians can be related to the evolution of prolonged gestation, which would not require prolonged external attachment of altricial young. Thus the occurrence of epipubic bones in two Cretaceous eutherians suggests that the dramatic modifications connected with typical placental reproduction^{3,6,7} may have been later events in the evolution of the Eutheria.

Since 1990, joint expeditions of the Mongolian Academy of Sciences and the American Museum of Natural History (MAE) have surveyed and collected the Cretaceous and Lower Tertiary sequences exposed in the Gobi Desert of Mongolia^{8,9}. In 1993 MAE discovered an extremely fossiliferous locality, Ukhaa Tolgod, noted for its Late Cretaceous assemblage of abundant and exquisitely preserved dinosaur skeletons, eggs and embryos, birds, lizards and mammals^{8–11}. More than 500 mammal skulls, many with well-preserved skeletons, have been collected at Ukhaa Tolgod. Skeletons of two eutherian taxa preserve epipubic bones: a zalmadalestid (cf. *Zalmadalestes*¹²), and a new taxon of insectivore-like eutherian closely related to the previously described *Asioryctes*¹³. The new taxon, *Ukhaatherium nessovi*, and its inclusive groups are here diagnosed:

Mammalia Linnaeus, 1758

Theria Parker and Haswell, 1897

Eutheria incertae sedis Gill, 1872

Asioryctitheria, new

Asioryctidae Kielan-Jaworowska, 1981

Ukhaatherium nessovi new genus and species; Figs 1, 2.

Etymology. Ukhaa (Mongolian), brown; theria (Latin), beast; nessovi, named after the late Lev Nessov for his pioneering work on Mesozoic mammals from Kazakhstan and Uzbekistan.

Age and locality. Late Cretaceous. Djadokhta? Formation, Ukhaa Tolgod, Mongolia.

Type and only species. *Ukhaatherium nessovi*.

Type specimen. PSS-MAE 102, skull with nearly complete skeleton.

Referred specimens. PSS-MAE 103–106, skulls and skeletons found in association with the type. PSS-MAE 110, skull with

lower jaws articulated, missing anterior snout and anterior mandible. PSS-MAE 111, complete skull and lower jaws articulated.

Diagnosis. *Ukhaatherium nessovi* bears a close resemblance to the Mongolian Late Cretaceous monotypic eutherians *Asioryctes* and *Kennalestes*, here united in the Asioryctitheria by the following characters: postglenoid vein exit within rather than posterior to postglenoid buttress, which is developed medially into an entoglenoid process; well-developed fusiform auditory bulla; pronounced caudal tympanic process of petromastoid (CTPP), connecting to promontorium by distinct interfenestral ridge; large piriform fenestra in anterior roof of tympanic cavity. *Ukhaatherium* and *Asioryctes* are grouped within Asioryctidae and separated from *Kennalestes* by the following characters¹³ P² (second upper premolar) smaller than P¹; upper molars more strongly elongated transversely, lacking pre- and postcingula; lower molars with smaller paraconids and more compressed trigonids, mastoid exposure rectangular in outline with large lower foramen; anterior contact of jugal with maxilla strongly bifurcate. *Ukhaatherium* differs from *Asioryctes* in having: enlarged, single-rooted upper canine (smaller and double-rooted in *Asioryctes*); less robust P³ with less salient paracone; large diastema between I⁵ (fifth upper incisor) and upper canine for occlusion of large lower canine; two mental foramina in lower jaw (between four and six in *Asioryctes*); smaller facial process of lacrimal so that nasal-maxillary contact in facial region is broader; and only one foramen in mastoid.

The asioryctitheres *Ukhaatherium*, *Asioryctes* and *Kennalestes* share features that indicate allocation to the Eutheria (Fig. 1) with close similarities to lipotyphlan insectivorans^{14,15}, including: a large maxilla component in the orbital wall and, probably, the concomitant early eruption of the cheek teeth¹⁶; an entoglenoid process^{14–16}; the features of the CTPP noted above; and a piriform fenestra.

These features notwithstanding, *Ukhaatherium* and its relatives *Kennalestes* and *Asioryctes* also have several primitive (plesiomorphic) skull and dental traits (for example, five upper incisors, posteriorly expansive nasals, facial process of the lacrimal, well-developed and posteriorly extensive jugal in the zygoma) radically

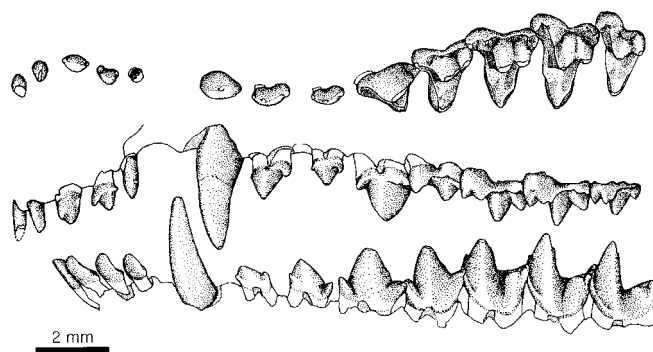


Figure 1 Dentition of the type specimen (PSS-MAE 102) of *Ukhaatherium nessovi*. Occlusal (top) and labial (middle) views of the left upper dentition: I^{1–5}, C, P^{1–4}, M^{1–3}. Labial view of lower dentition (bottom): I_{1–4}, C, P_{1–4}, M_{1–3}. Lower dentition is a composite of the right I_{1–3} and C, and the left I₄, P_{1–4} and M_{1–3}. The dentition of *U. nessovi* shows several derived traits shared with Placentalia (here defined as the clade composed of the most recent common ancestor of living eutherians and its descendants): three molars, reduced styler shelf, small paraconid, compressed trigonid, complex posterior premolars, and other putative distinctive features, such as paracone larger than metacone and absence of styler cusp C. Other cranial diagnostic features include the presence of an optic foramen, a small facial process of the lacrimal, a slender zygomatic arch, presence of ectopterygoid flanges of the alisphenoid, and a well-developed crista interfenestralis connecting the promontorium with the paroccipital process. Primitive features absent among placentalia, but present in *U. nessovi* and closely allied forms, suggest these Late Cretaceous forms from Mongolia are basal members of Eutheria (mammals more closely related to Placentalia than to Metatheria).

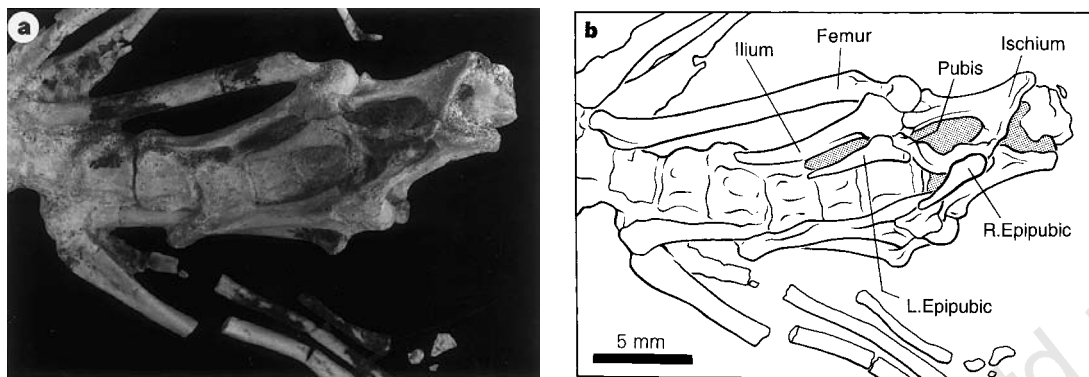


Figure 2 Detail of the skeleton (PSS-MAE 105) of *Ukhaatherium nessovi* showing the pelvic girdle and epipubic elements. Ventral view. **a**, Photograph; **b**, labelled sketch. The epipubics are posteriorly expanded and plate-like near the point of articulation with the pubis. They taper anteriorly in a splint-like process with a

slight concave upward curvature. A roughened ridge on the anterolateral corner of the pubis indicates the point of contact with the epipubics. Epipubic bones are not as well preserved in the type specimen PSS-MAE 102, but they are very similar in proportion and shape to those of the skeleton PSS-MAE 105. L., left; R., right.

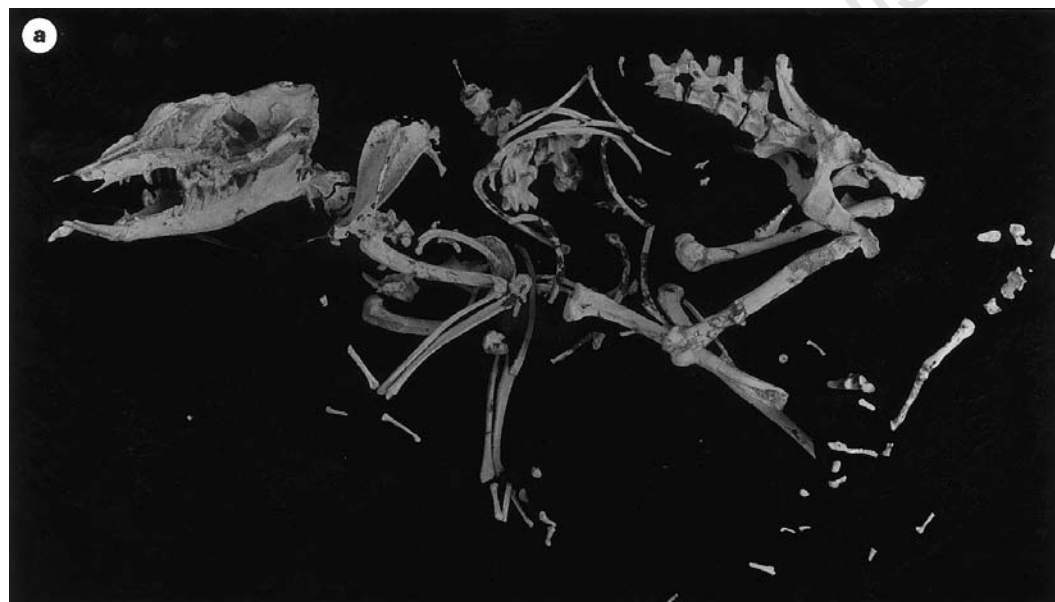
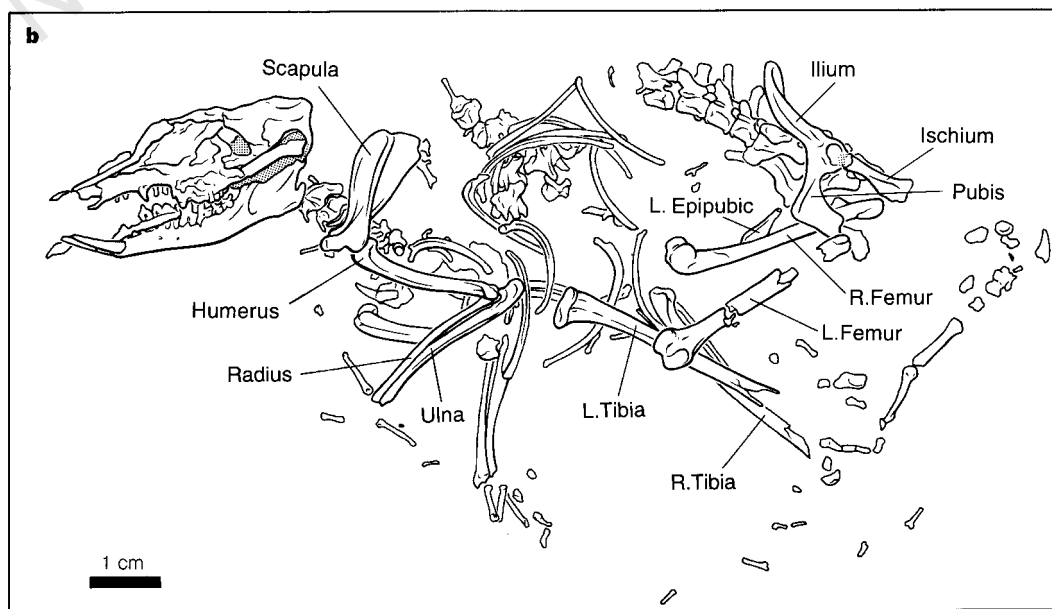


Figure 3 Lateral view of the skeleton (PSS-MAE 131) of cf. *Zalambdalestes*. **a**, Photograph; **b**, labelled sketch. L., left; R., right.



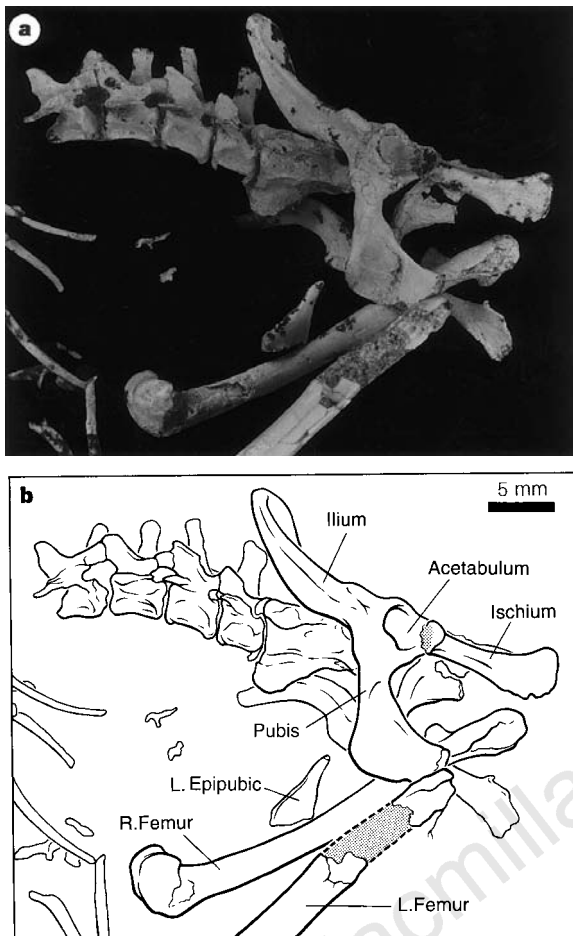


Figure 4 Detail of the skeleton (PSS-MAE 131) of cf. *Zalambdalestes* showing pelvic girdle and epipubic bone. **a**, Photograph; **b**, labelled sketch. The epipubic has a large, plate-like posterior moiety that tapers anteriorly, but the anterior section has been broken off. L., left; R., right.

modified in lipotyphlans and many other eutherian groups. Unlike lipotyphlans, Asioryctitheria (where relevant skeletal evidence is known) do not show a marked reduction or loss of the pubic symphysis¹⁶. Moreover, postcranial skeletons of *Asioryctes* retain primitive characters in pedal and vertebral structure markedly modified in insectivorans and other eutherian mammals¹⁷.

The presence of epipubic bones in *Ukhaatherium nessovi* is also a primitive aspect of this taxon (Fig. 2). An epipubic bone is also preserved in a skeleton of cf. *Zalambdalestes* (Figs 3, 4). This substantiates Kielan-Jaworowska's⁷ prediction of epipubic bones in zalambdalestids, which was based on a triangular fossa on the anteroventral corner of the pubis in the related *Barunlestes butleri*.

Zalambdalestids lack many of the above-mentioned diagnostic cranial features of the Asioryctitheria and have a specialized dentition (including enlarged procumbent incisors) and skeleton (elongated limbs, fusion of the tibia-fibula, and restricted ankle movement) that are reminiscent of features found in rabbits. Although the affinities of zalambdalestids are still a matter of debate^{14,18}, their clear distinction from the Asioryctitheria suggests early differentiation of eutherian groups that retain epipubic bones and other primitive mammalian skeletal traits.

Epipubic bones occur in monotremes, marsupials, Mesozoic mammals (multituberculates, a symmetrodont (Y. Hu, Y. Wang, C.-K. Li & Z. Luo, manuscript in preparation) and cladotheres^{19–22}), and allied tritylodonts, suggesting that their presence is the primitive mammalian condition. Indeed, epipubic bones have been

homologized with the anterior part of the puboischiadic plate in *Sphenodon*, lizards and therapsids, as well as with the baculum of the male and the os clitoridis of the female eutherian²³. It has been argued that, in female marsupials, epipubic bones support the pouch and pouch young^{1,3,4}. This has been disputed because epipubic bones in marsupials are found in males and in pouchless females^{24–26}. Nonetheless, epipubic bones show sexual dimorphism (they are either longer or broader in females than in males of the same species), and probably function in both locomotion and reproduction⁵. These elements could support either a pouch^{1,3} or the teats and the suspended young in pouchless marsupials⁵. Loss of epipubic bones is widely acknowledged to be a derived diagnostic trait of eutherians^{14,24}, possibly associated with the origin of prolonged intrauterine development^{3,6,7} that obviated the need for prolonged support of the externally attached young. The abdominal rigidity conferred by epipubic bones would also be disadvantageous in placental mammals, where prolonged gestation is associated with abdominal expansion during pregnancy³.

The occurrence together and ancestral status of epipubic bones and the prolonged external attachment of young does not of course demonstrate that these traits are functionally dependent on one another, and the problem requires further functional and anatomical study. However, the possibility of such functional interdependence has been recognized⁵. Thus the Mongolian Cretaceous eutherians, unlike placentals (living eutherians), might have had a short gestation period that necessitated prolonged suspension of the altricial young to the teats before weaning. A pouch might also have been present, but even its ancestral status in marsupials is disputed²⁷ or equivocal²⁸.

Another feature that has been related to reproductive modes is the acute V-shaped angle of the ischial arc that is seen in both *Ukhaatherium* and zalambdalestids⁷. The acute ischial arc is found in many living marsupials and eutherians with smaller and less-developed offspring. Because the ischial arc tends to be broad in viviparous mammals with extended gestation, it has been argued that the acute arc in combination with the epipubic bones in fossil mammals may indicate a short gestation time⁷.

Regardless of its implications for reproductive evolution, evidence that Asioryctitheria and Zalambdalestidae retain epipubics refutes a traditional diagnostic feature of the Eutheria. This and other primitive features of Asioryctitheria suggest a basal position for this group among eutherians. Further corroboration of this pattern would provide a useful framework for the difficult problem of higher eutherian phylogeny²⁹, where the basal branches are, at present, poorly resolved. At a more inferential level, this pattern raises the possibility that the innovations in the reproductive system diagnosing the crown-group Placentalia^{6,25} may not have been present in early members of Eutheria. □

Received 26 June; accepted 8 September 1997.

1. Leche, W. Zur Morphologie der Beutelknochen. *Biol. Fören. Förhandl. (Stockh.)* **3**, 120–126 (1891).
2. Fourie, S. Notes on a new tritylodontid from the Cave Sandstone of South Africa. *Navors. Nasion. Mus.* **2**, 7–19 (1962).
3. Lillegraven, J. A. Latest Cretaceous mammals of upper part of Edmonton Formation of Alberta, Canada, and review of marsupial-placental dichotomy in mammalian evolution. *Univ. Kansas Paleontol. Contrib.* **50**, 1–122 (1969).
4. Vaughn, P. P. The phylogenetic migrations of the ambiens muscle. *J. Elisha Mitchell Sci. Soc.* **72**, 243–262 (1956).
5. White, T. D. An analysis of epipubic bone function in mammals using scaling theory. *J. Theor. Biol.* **139**, 343–357 (1989).
6. Lillegraven, J. A., Thompson, S. D., McNab, B. K. & Patton, J. L. The origin of eutherian mammals. *Biol. J. Linn. Soc.* **32**, 281–336 (1987).
7. Kielan-Jaworowska, Z. Possible occurrence of marsupial bones in Cretaceous eutherian mammals. *Nature* **255**, 698–699 (1975).
8. Novacek, M. J., Norell, M., McKenna, M. C. & Clark, J. Fossils of the Flaming Cliffs. *Sci. Am.* **271**, 60–69 (1994).
9. Dashzeveg, D. et al. Extraordinary preservation in a new vertebrate assemblage from the Late Cretaceous of Mongolia. *Nature* **374**, 446–449 (1995).
10. Norell, M. A. et al. A theropod dinosaur embryo and the affinities of the Flaming Cliffs dinosaur eggs. *Science* **266**, 779–782 (1994).
11. Norell, M. A., Clark, J. M., Chiappe, L. M. & Dashzeveg, D. A nesting dinosaur. *Nature* **378**, 774–776 (1995).
12. Gregory, W. K. & Simpson, G. G. Cretaceous mammal skulls from Mongolia. *Am. Mus. Novit.* **225**, 1–20 (1926).

13. Kielan-Jaworowska, Z. Preliminary description of two new eutherian genera from the Late Cretaceous of Mongolia. *Palaeontol. Pol.* **33**, 5–16 (1975).
14. Novacek, M. J. The skull of lepidictid insectivores and the higher-level classification of eutherian mammals. *Bull. Am. Mus. Nat. Hist.* **183**, 1–111 (1986).
15. Kielan-Jaworowska, Z. Evolution of the therian mammals in the Late Cretaceous of Asia. Part IV. Skull structure in *Kennalestes* and *Asioryctes*. *Palaeontol. Pol.* **42**, 25–78 (1981).
16. MacPhee, R. D. E. & Novacek, M. J. in *Mammal Phylogeny, Placentals* (eds Szalay, F. S., Novacek, M. J. & McKenna, M. C.) 13–31 (Springer, New York, 1993).
17. Kielan-Jaworowska, Z. Evolution of the therian mammals in the Late Cretaceous of Asia. Part II. Postcranial skeleton in *Kennalestes* and *Asioryctes*. *Palaeontol. Pol.* **37**, 55–83 (1977).
18. McKenna, M. C. in *Phylogeny of the Primates: A Multidisciplinary Approach* (eds Luckett, W. P. & Szalay, F. S.) 21–46 (Plenum, New York, 1975).
19. Krause, D. W. & Jenkins, F. A. Jr The postcranial skeleton of North American multituberculates. *Bull. Mus. Comp. Zool.* **150**, 199–246 (1983).
20. Kielan-Jaworowska, Z. & Gambaryan, P. P. Postcranial anatomy and habits of Asian multituberculate mammals. *Fossils Strata* **36**, 1–92 (1994).
21. Krebs, B. Das Skelett von *Henkelotherium guimorotae* gen. et sp. nov. (Eupantotheria, Mammalia) aus dem Oberen Jura von Portugal. *Berl. Geowiss. Abh.* **A 133**, 1–121 (1991).
22. Rougier, G. W. *Vincelestes neuquenianus* Bonaparte (Mammalia, Theria). Un primitivo mamífero del Cretácico Inferior de la Cuenca Neuquina. Dissertation, Univ. Buenos Aires (1993).
23. Jellison, W. L. A suggested homolog of the os penis or baculum of mammals. *J. Mammal.* **26**, 146–147 (1945).
24. Marshall, L. G. Evolution of the metatherian and eutherian (mammalian) characters: a review based on cladistic methodology. *Zool. J. Linn. Soc.* **66**, 369–410 (1979).
25. Tyndale-Biscoe, C. H. & Renfree, M. B. *Reproductive Physiology of Marsupials* (Cambridge Univ. Press, 1987).
26. Renfree, M. B. in *Mammal Phylogeny, Mesozoic Differentiation, Multituberculates Monotremes, Early Therians and Marsupials* (eds Szalay, F. S., Novacek, M. J. & McKenna, M. C.) 4–20 (Springer, New York, 1993).
27. Reig, O. A., Kirsch, J. A. W. & Marshall, L. G. in *Possums and Opossums: Studies in Evolution* (ed. Archer, M.) 1–90 (Surrey Beatty, Chipping Norton, 1987).
28. Springer, M. S., Kirsch, J. A. W. & Case, J. A. in *Molecular Evolution and Adaptive Radiations* (eds Givnish, T. J. & Sytsma, K.) 129–161 (Cambridge Univ. Press, New York, 1997).
29. Novacek, M. J. Mammalian phylogeny: shaking the tree. *Nature* **356**, 121–125 (1992).

Acknowledgements. We thank A. Davidson for preparation of the fossils and E. Heck for the illustrations. This work was supported by the National Science Foundation, the Jaffe Foundation, the National Geographic Society, the Eppley Foundation, and the James Carter Memorial Fund and the Frick Laboratory Endowment of The American Museum of Natural History.

Correspondence and requests for materials should be addressed to M.J.N. (e-mail: novacek@amnh.org).

The origin of the dog-like borhyaenoid marsupials of South America

Christian de Muizon*, Richard L. Cifelli† & Ricardo Céspedes Paz‡

* URA 12 CNRS, Muséum National d'Histoire Naturelle, Laboratoire de Paléontologie, 8 rue Buffon, Paris F-75005, France

† Oklahoma Museum of Natural History, University of Oklahoma, Norman, Oklahoma 73019, USA

‡ Universidad Mayor San Simón, Museo Arqueológico, Cas. 992, Cochabamba, Bolivia

Dog-like marsupials (superfamily Borhyaenoidea) were the largest predacious mammals during the Tertiary period in South America¹. They are critical to our understanding of marsupial origin, phylogeny and palaeobiogeography because they have been related to various marsupial lineages of several continents: didelphoids² (mainly New World, but also Europe, Asia and Africa), pelyomyids³, stagodontids⁴ (North America), dasyuroids⁵ (Australia) and deltatheroidans⁶ (predominantly Asian). These relationships, based mainly on dental morphology, have been discussed and rejected several times^{2,3,7,8}. Here we report the discovery of exceptionally well preserved skulls and skeletons, referable to the didelphoid *Andinodelphys*, which shed new light on the phylogenetic and palaeobiogeographic origin of dog-like marsupials. The skulls of *Mayulestes* (borhyaenoid)⁹, *Andinodelphys* and *Pucadelphys* (didelphoids)^{10,11} from the early Palaeocene epoch of Bolivia are the oldest known for American marsupials. Comparison of their basicranial anatomy suggests that dog-like marsupials are closely related to an early didelphimorphian radiation in South America, rather than to Asiatic (deltatheroidan), North American (stagodontid), or Australian (dasyuroid) lineages.

The skulls of *Andinodelphys* (Fig. 1), *Pucadelphys*¹⁰ and *Mayulestes*⁹ bear a medial process of the squamosal^{9,12}, an apomorphic character state unknown in other marsupials and regarded here as a synapomorphy of these three genera. They lack a tympanic process of the alisphenoid and have a foramen ovale limited posteriorly by the periotic, two plesiomorphic conditions, the former being unknown in other marsupials^{9,12} (Fig. 2). *Mayulestes* is more derived than *Pucadelphys* and *Andinodelphys* in having a small alisphenoid hypotympanic sinus and in lacking a prootic canal, two borhyaenoid synapomorphies^{9,12}. *Mayulestes* exhibits a complex of dental adaptations to a hypercarnivorous diet, independently evolved in several groups of mammals^{2,7,9}. The presence of a large nasolacrimal contact in *Mayulestes* is likely to be a plesiomorphic feature because it is found in other borhyaenoids^{10,12}, cynodonts, non-therian mammals, and *Vincelestes*. Because this feature is occasionally found in some Australian marsupials (some wynyardiids, vombatids and phalangerids)², its loss, which seems to be widespread in marsupials, probably occurred several times independently in various lineages. *Andinodelphys*, *Pucadelphys* and all the other didelphoids have a frontomaxillary contact that precludes a nasolacrimal contact. The three Tiupampian genera also share an enlarged and anteriorly projected first upper incisor, a staggered second lower incisor, large pterygoid hamuli, a large contribution of the pars mastoidea of periotic to the occiput. *Andinodelphys* differs from *Pucadelphys* and *Mayulestes* in having small palatal fenestrae and a small transverse canal.

Molar structure, which is commonly used in the interpretation of marsupial systematics⁴, relates both *Pucadelphys* and *Andinodelphys* (Fig. 1) to the didelphoid didelphimorphians¹³, as they have the paracone smaller than the metacone, incipiently V-shaped centrocrista, a robust protocone with conules and swollen posterior base, paraconid lower than metaconid, and loss of labially extended postprotocrista, a well-developed styler shelf with large cusps B and D, styler cusp C small, sometimes twinned, although these features may be ambiguous¹⁴. Recent didelphoids do not have a medial process of the squamosal, so, either this feature is a didelphimorphian synapomorphy that disappeared in the lineage leading to Recent taxa, or *Mayulestes*, *Pucadelphys* and *Andinodelphys* are part of a monophyletic subgroup of the Didelphimorphia that acquired this synapomorphy.

Combined craniodental evidence of the three well-represented taxa from Tiupampa suggests that borhyaenoids originated from a plesiomorphic didelphimorphian stem group (Fig. 3), from a species with the following features: a medial process of the squamosal (a possible synapomorphy of the Didelphimorphia, but lost in some lineages), no tympanic process of the alisphenoid, no alisphenoid hypotympanic sinus, a pro-otic canal, a large foramen ovale limited by the alisphenoid anteriorly and by the periotic posteriorly, no transverse canal, no palatal vacuities, a conspicuous nasolacrimal contact, a paracone slightly smaller than (or subequal to) the metacone, a straight centrocrista (the acquisition of a V-shaped centrocrista is a didelphoid synapomorphy, although it also occurs in most dasyuroids), a large styler shelf with well-developed styler cusps B and D, and a postprotocrista that does not extend labially past the base of the metacone. The presence of styler cusp C in this hypothetical ancestor is debatable, as this feature may have appeared and disappeared several times in marsupial evolution^{15–17}. However, given the great variability of the size of styler cusp C in *Andinodelphys*, a primitive didelphimorphian, we suggest that early didelphimorphians had either a very small or no styler cusp C.

The phylogenetic relationships shown in Fig. 3 imply that some basicranial character states (tympanic process of the alisphenoid, middle-ear sinuses) probably appeared several times in marsupial evolution⁹. Furthermore, among the abundant dental characters that are generally considered in marsupial phylogenies, we retain a few we consider least equivocal. The size of the protocone, V-shaped centrocrista, styler cusp C, the width of the talonid, and the dental

functional complex related to carnivorous diet are regarded here as homoplastic features that are very likely to have appeared (and/or disappeared) several times in marsupial evolution.

This interpretation (Fig. 3) does not support a close relationship between borhyaenoids and deltatheroidans⁶, with similarities between the two groups being regarded here as symplesiomorphies or homoplasies, as discussed elsewhere¹². Furthermore, it is commonly accepted^{8,12,17} that stagodontids are unlikely to have close relationships with borhyaenoids⁴. This interpretation is confirmed by the morphology of the ear region of stagodontids, which is more specialized than that of *Mayulestes* in possessing a large recess for the external auditory meatus and a tympanic process of the alisphenoid, and which, apparently, lacks a medial process of the squamosal^{18,19}. Furthermore, there are important pedal similarities between borhyaenoids and caenolestoids¹⁷, although the two groups have a fairly different dental and cranial anatomy. The fact that no significant caenolestoid cranial remains are known in the early Tertiary prevents any comparison with the Tiupampa marsupials.

It is commonly believed that marsupials from South America had their origin in a Holarctic marsupial stock probably derived from North America^{13,20,21}. Among Late Cretaceous taxa from North America (represented by dental specimens only), some species of *Protalphadon* (*P. lulli*, *P. waheapensis*) and *Anchistodelphys*¹⁶ are dentally compatible with an ancestral position to the Didelphimorphia. However, because these taxa bear a clear extension of the postprotocrista below the posterior edge of the metacone (a feature present in more distal taxa such as *Kokopellia*^{15,22} and regarded as a

plesiomorphic character state for marsupials²³), the absence of this feature in the Didelphimorphia must be regarded as a synapomorphy of the group. Nevertheless, dental features seem to be so homoplastic and variable^{1,4,15,22–26} that identification of the ancestral group for the Didelphimorphia will remain highly conjectural until good cranial remains are found in North America. The South American microbiotheres have been regarded as the sister-group of all the Australian marsupials¹⁷ on the basis of their continuous lower ankle joint. Because of the lack of significant cranial remains in the early Tertiary, they could not be compared to the Tiupampa Didelphimorphia, and node 7 of Fig. 3 (which implies the loss of synapomorphies of node 5 in microbiotheres) has to be tested by the discovery of good cranial remains in the early Tertiary. The same is to be said about the possible close affinities of microbiotheres and pardiomyids, which are based only on the reduction of the styler shelf and its cusps⁴.

The timing of arrival of higher mammals in South America has been abundantly discussed^{13,20,27,28}. In this context, the Tiupampa assemblage, which is the oldest well-represented, diverse fauna on the continent, is particularly important. The affinities of the Tiupampa mammals clearly suggest a diachronous arrival. Placental mammals are highly similar to those of the North American early Palaeocene, suggesting a relatively recent common ancestry^{13,26} and, hence, recent dispersal between North and South America. The marsupial fauna, however, is considerably more diverse and more divergent from North American assemblages^{13,26}. The Tiupampa didelphimorphians apparently have closer relationships to other

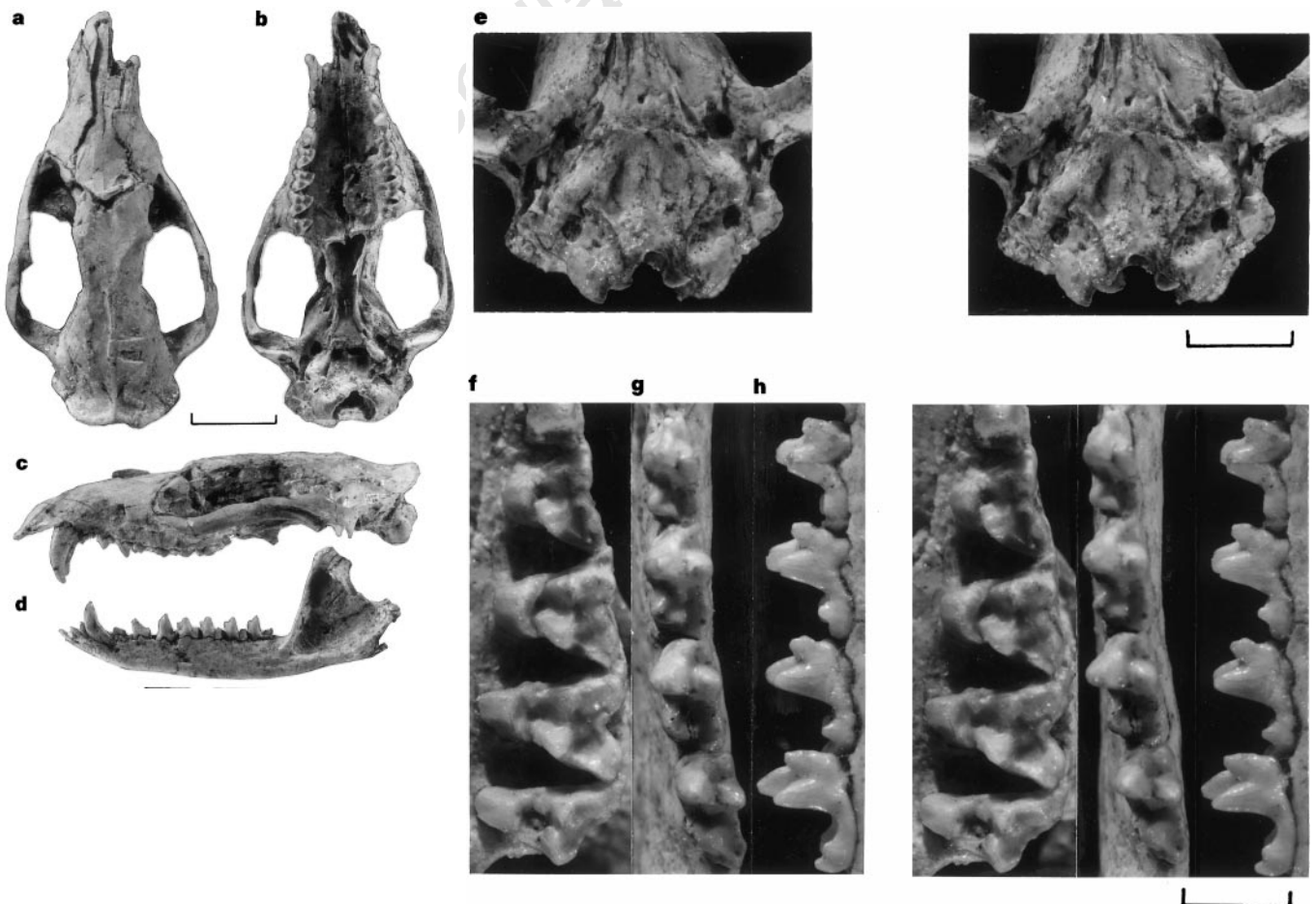


Figure 1 Views of *Andinodelphys cochabambensis* skull and teeth. **a–c**, Dorsal (**a**), ventral (**b**) and lateral (**c**) views of the *A. cochabambensis* skull; **d**, left hemimandible in lateral view; **e**, basicranium; **f**, upper molars and premolars in

occlusal view; **g**, **h**, lower molars: occlusal (**g**) and medial (**h**) views. **a–d**, specimen MHNC 8264; **e**, **g**, **h**, MHNC 8308; **f**, MHNC 8364. Scale bars; **a–d**, 1 cm; **e**, 5 mm; **f–h**, 3 mm.

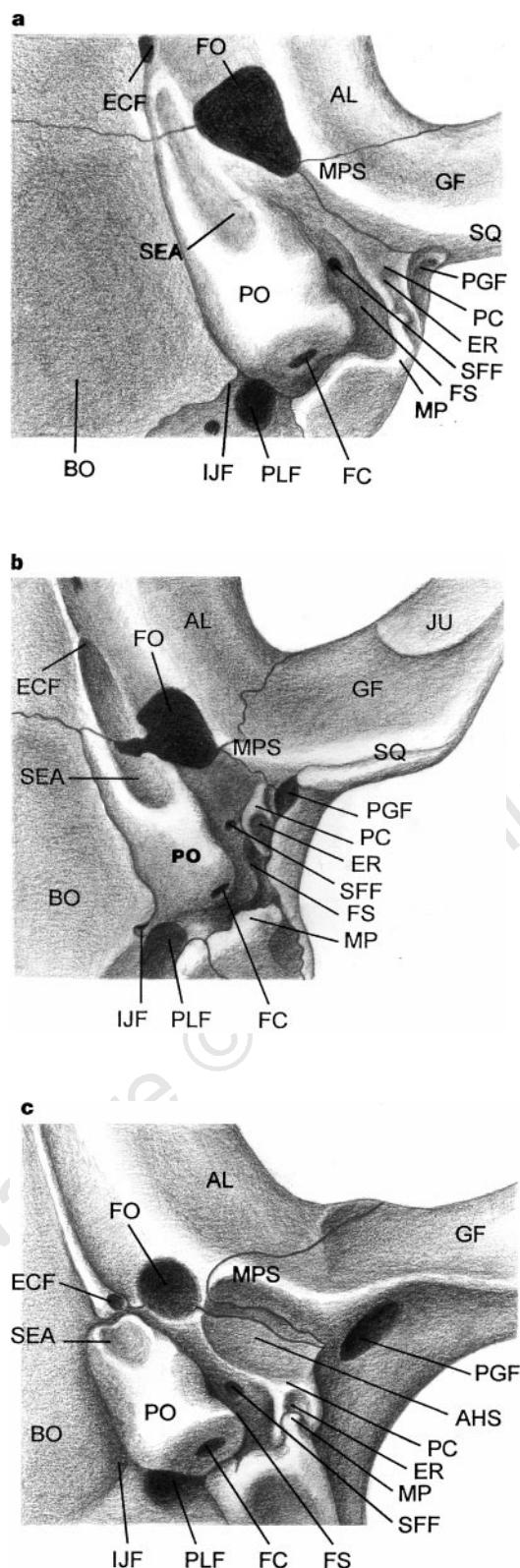


Figure 2 Comparative reconstructions of the ear regions of *Pucadelphys*, *Andinodelphys* and *Mayulestes*. **a**, *Pucadelphys*; **b**, *Andinodelphys*; **c**, *Mayulestes*. Abbreviations: AHS, alisphenoid hypotympanic sinus; AL, alisphenoid; BO, basioccipital; ECF, entocarotid foramen; ER, epitympanic recess; FC, fenestra cochlearis; FO, foramen ovale; FS, facial sulcus; GF, glenoid fossa; IJF, inferior jugular foramen; MP, mastoid process; MPS, medial process of the squamosal; PC, petrosal crest; PGF, postglenoid foramen; PLF, posterior lacerate foramen; PO, promontorium; SEA, sulcus for the entocarotid artery; SFF, secondary facial foramen; SQ, squamosal.

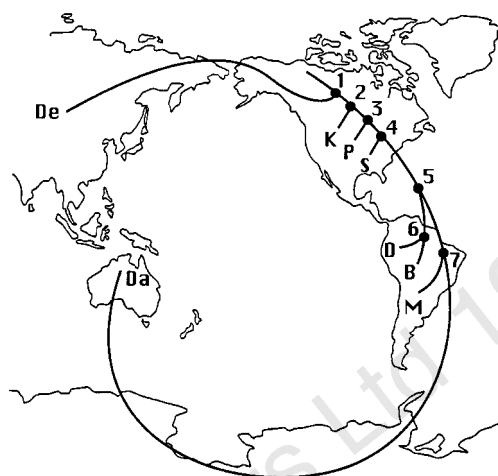


Figure 3 Distribution and phylogenetic relationships of the major groups which have been closely related to dog-like marsupials. Also included are the Albian/Cenomanian North American genus *Kokopellia* (which represents a basal morphology for marsupials¹⁵) and the microbiotheres. Node 1 (Metatheria): number of premolars reduced to three. Node 2: postprotocrista extends on base of metacone, and presence of labial postcingulid. Node 3: acquisition of a stylar cusp D, and twinning of entoconid and hypoconulid. Node 4: loss of labial extension of postprotocrista. Node 5: staggered i2 (this character state is absent in *Eodelphis*²⁹, *Kokopellia*²² and *Alphadon*²⁹) and increase in size of I1, which protrudes anteriorly (this character state tends to disappear in some borhyaenoids, thylacynids, some dasyurids, and microbiotheres). Node 6 (Didelphimorphia): medial process of the squamosal. Node 7: large tympanic process of the periotic, continuous lower ankle joint¹⁷, sperm structure, albumin data³⁰. Abbreviations: B, borhyaenoids; Da, dasyuromorphs; De, deltatheroidans; D, didelphids; K, *Kokopellia*; P, pemiomyids; S, stagodontids.

South American than North American taxa, corroborating previous suggestions of greater antiquity for the dispersal of marsupials into South America^{13,26}. □

Received 24 June; accepted 5 September 1997.

- Marshall, L. G. Evolution of the Borhyaenidae, extinct south American predaceous marsupials. *Univ. Calif. Pub. Geol. Sci.* **117**, 1–89 (1978).
- Archer, M. in *Carnivorous Marsupials* (ed. Archer, M.) 397–443 (Royal Society of New South Wales, Sydney, 1982).
- Marshall, L. G. Cladistic analysis of borhyaenoid, dasyuroid, didelphoid, and thylacynid (Marsupialia, Mammalia) affinity. *Syst. Zool.* **26**, 410–425 (1977).
- Marshall, L. G., Case, J. A. & Woodburne, M. O. in *Current Mammalogy* (ed. Genoways, H. H.) 343–505 (Plenum, New York, 1990).
- Archer, M. The dasyurid dentition and its relationships to that of didelphids, thylacynids, borhyaenids (Marsupiacarnivora) and peramelids (Peramelina, Marsupialia). *Aust. J. Zool. (suppl.)* **39**, 1–34 (1976).
- Marshall, L. G. & Kielan-Jaworowska, Z. Relationships of the dog-like marsupials, deltatheroidans, and early tribosphenic mammals. *Lethaia* **25**, 361–374 (1992).
- Fox, R. C. Ancestry of the “dog-like” marsupials. *J. Paleontol.* **53**, 733–735 (1979).
- Fox, R. C. & Naylor, B. G. in *Sixth Symp. Mesozoic Terrestrial Ecosystem Biota* (eds San, A. & Wang, Y.) 247–250 (China Ocean, Beijing, 1995).
- Muizon, C. de A new carnivorous marsupial from the early Palaeocene of Bolivia and the problem of marsupial monophyly. *Nature* **370**, 208–211 (1994).
- Marshall, L. G. & Muizon, C. de *Mém. Mus. Natl. Hist. Nat.* **165**, 21–90 (1995).
- Marshall, L. G. & Muizon, C. de The dawn of the age of mammals in South America. *Natl. Geogr. Res.* **4**, 23–55 (1988).
- Muizon, C. de *Mayulestes ferox*, a borhyaenoid (Metatheria, Mammalia) from the early Palaeocene of Bolivia. Phylogenetic and Palaeobiologic implications. *Géodiversitas* (in the press).
- Muizon, C. de *Rev. Técn. Yac. Petrol. Fisc. Bol.* **12**, 393–433 (1992).
- Cifelli, R. L. Cretaceous mammals from southern Utah. I. Marsupials from the Kaiparowits Formation (Judithian). *J. Vert. Paleontol.* **10**, 295–319 (1990).
- Cifelli, R. L. Early Cretaceous mammal from North America and the evolution of marsupial dental characters. *Proc. Natl. Acad. Sci. USA* **90**, 9413–9416 (1993).
- Cifelli, R. L. in *Mammal Phylogeny* (eds Szalay, F. S., Novacek, M. J. & McKenna, M. C.) 205–215 (Springer, New York, 1993).
- Szalay, F. S. *Evolutionary History of the Marsupials and an Analysis of Osteological Characters* (Cambridge Univ. Press, 1994).
- Clemens, W. A. Fossil mammals of the type Lance Formation, Wyoming. Part II. Marsupialia. *Univ. Calif. Publ. Geol. Sci.* **62**, 1–122 (1966).
- Matthew, W. D. A marsupial from the Belly River Cretaceous. With critical observations upon the affinities of the Cretaceous mammals. *Bull. Am. Mus. Nat. Hist.* **35**, 447–500 (1916).
- Hoffstetter, R. Données et hypothèses concernant l'origine et l'histoire biogéographique des marsupiaux *C. r. Acad. Sci.* **274**, 2635–2638 (1972).

21. Marshall, L. G. in *Aspects of Vertebrate Evolution* (ed. Jacobs, L. L.) 345–386 (Museum of Northern Arizona Press, Flagstaff, 1980).
22. Cifelli, R. L. & Muizon, C. de Dentition and Jaw of *Kokopellia juddi*, a primitive marsupial or near-marsupial from the medial Cretaceous of Utah. *J. Mamm. Evol.* (in the press).
23. Clemens, W. A. & Lillegraven, J. A. New Late Cretaceous North American advanced therians that fit neither the marsupial nor the therian molds. *Contrib. Geol. Spec. Pap.* 3, 55–85 (1986).
24. Szalay, F. S. & Trofimov, B. The Mongolian Late Cretaceous *Asiatherium*, and the early phylogeny and paleobiogeography of Metatheria. *J. Vert. Paleontol.* 16, 474–509 (1996).
25. Fox, R. C. in *Possums and Opossums* (ed. Archer, M.) 161–169 (Surrey Beatty/Royal Society of New South Wales, Sydney, 1987).
26. Muizon, C. de & Brito, I. Le bassin calcaire de São José de Itaboraí (Rio de Janeiro, Brésil): ses relations avec le site de Tiupampa (Cochabamba, Bolivie). *Ann. Paleontol.* 79, 233–269 (1993).
27. Simpson, G. G. *Splendid Isolation. The Curious History of South American Mammals* (Yale Univ. Press, New Haven, CT, 1980).
28. Bonaparte, J. New Late Cretaceous mammals from the Los Alamos Formation, northern Patagonia. *Natl Geogr. Res.* 6, 63–94 (1990).
29. Cifelli, R. L. & Muizon, C. de Marsupial mammals from the upper Cretaceous North Horn Formation, central Utah. *J. Paleontol.* (in the press).
30. Aplin, K. P. & Archer, M. in *Possums and Opossums* (ed. Archer, M.) XV–LXXII (Surrey Beatty/Royal Society of New South Wales, Sydney, 1987).

Acknowledgements. We thank R. Suarez Soruco for logistic support; M. Archer, Z. Luo and M. Woodburne for comments; F. Pilar for Fig. 2; H. Lavina for Fig. 3; and D. Serrette for photography. This work was funded by the Institut Français d'Études Andines (Lima, Peru), the URA 12 of the Centre National de la Recherche Scientifique (Paris, France), and by the Museum National d'Histoire naturelle (Paris, France).

Correspondence and requests for materials should be addressed to C.d.M. or R.L.C.

The first skull of *Australopithecus boisei*

Gen Suwa*, Berhane Asfaw†, Yonas Beyene‡, Tim D. White§, Shigehiro Katoh||, Shinji Nagaoka¶, Hideo Nakaya#, Kazuhiro Uzawa*, Paul Renne** & Giday WoldeGabriel**

* Department of Biological Sciences, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan

† Rift Valley Research Service, PO Box 5717, Addis Ababa, Ethiopia

‡ Department of Archaeology and Anthropology, CRCC, Ministry of Information and Culture, PO Box 1907, Addis Ababa, Ethiopia

§ Laboratory for Human Evolutionary Studies, University of California, Berkeley, California 94720, USA

|| Division of Earth Sciences, Hyogo Museum of Human and Nature Activities, Yayoigaoka, Sanda-shi, Hyogo 669-13, Japan

¶ Department of Geography, Nagasaki University, Bunkyo-machi, Nagasaki-shi, Nagasaki 852, Japan

Department of Earth Sciences, Kagawa University, Saiwai-cho, Takamatsu-shi, Kagawa 760, Japan

* Berkeley Geochronology Center, 2455 Ridge Road, Berkeley, California 94709, USA

** EES-1/D462, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

Australopithecus boisei was first described from a cranium recovered in 1959 from Olduvai Gorge, Tanzania^{1,2}. This and subsequent finds, mostly from Kenya's Turkana basin^{3–5}, resulted in its characterization as a specialized *Australopithecus* species with a hyper-robust masticatory apparatus^{2,4,6}. A distinct *A. boisei* facial morphology has been emphasized to differentiate robust *Australopithecus* lineages from East and South Africa⁶. A preference for closed and/or wet habitats has been hypothesized⁷. Here we report some new *A. boisei* specimens, including the taxon's first cranium and associated mandible, from Konso, Ethiopia. These fossils extend the known geographical range of *A. boisei*. They provide clear evidence for the coexistence of *A. boisei* and *Homo erectus* within a predominantly dry grassland environment. The *A. boisei* specimens from Konso demonstrate considerable morphological variation within the species. The unexpected combination of cranial and facial features of this skull cautions against the excessive taxonomic splitting of early hominids based on morphological detail documented in small and/or geographically restricted samples.

The fossiliferous deposits of Konso (initially Konso–Gardula) were discovered in 1991 (ref. 8). Remains of *H. erectus* and the oldest firmly dated Acheulean assemblages were reported. Fieldwork since 1993 has clarified the temporal and spatial distributions of Konso's fossiliferous sequence. The early Pleistocene series comprises discontinuous outcrops extending for about 15 km. Field relations of marker beds, tephrostratigraphy and ⁴⁰Ar/³⁹Ar dating of tuffs have established that the bulk of the Konso fossils derive from two time horizons, one at approximately 1.9 million years ago (Myr) and the other at ~1.4 Myr. Less-extensive collections have been made between these levels.

The ~1.4-Myr fossiliferous strata span a composite section of approximately 50 m, extending discontinuously for about 10 km. All *A. boisei* specimens were recovered at one locality, KGA10, where alternating clay and silt/sand layers are exposed for ~1 km. The highly fossiliferous sands at the mid-section of KGA10 are interpreted to be the middle to distal portions of an alluvial fan, deposited adjacent to, and extending into, a lake. Fossils and artefacts deriving from horizons of sands and silts are not abraded and show evidence of minimal transport. A large mammalian assemblage has been collected from these deposits, showing a striking dominance of Alcelaphini (80% of over 800 identified bovid craniodental remains), and common occurrence of equids and two *Metridiochoerus* species. We interpret this to indicate the presence of extensive dry grasslands at KGA10.

Mammalian fauna from other localities/horizons at Konso suggest the dominance of more wet and/or closed environments. These include seven moderate to large fossil assemblages (consisting of over 100 to 1,800 identifiable mammalian specimens, excluding hippopotamuses), with Reduncini and Bovini making up over 50% to 70% of bovids in each assemblage. Alcelaphini remains are fewer, its relative abundance within bovids ranging between approximately 10% and 40%. These assemblages derive from limited horizons of sand layers and lenses occurring within sections dominated by finer clays and silts, and are interpreted to be indicative of mesic riverine habitats or edaphic grasslands at lake margin settings.

All nine Konso *A. boisei* specimens were recovered among the predominantly dry grassland fauna of KGA10 (consisting of over

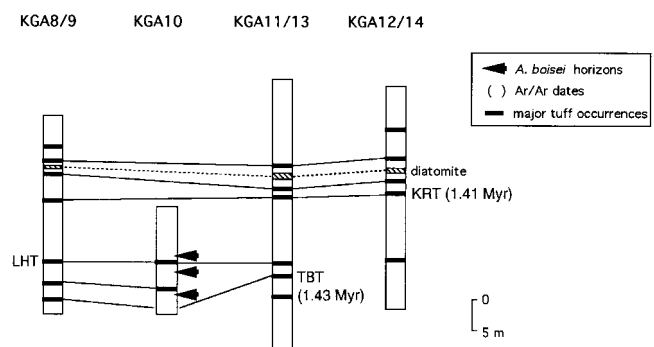


Figure 1 A schematic stratigraphy of the Konso deposits bearing *A. boisei*. Unbroken lines indicate tuff correlations based on lithostratigraphy and major-element compositions of single glass shards. The ⁴⁰Ar/³⁹Ar analysis followed established procedures using laser fusion of single crystals³⁰. The Karat tuff (KRT) was dated from 18 sanidine crystals with a weighted mean of 1.413 Myr (standard error 0.023 Myr). The Trail Bottom tuff (TBT) was dated from 20 sanidine crystals with a weighted mean of 1.430 Myr (standard error 0.023 Myr). The Lehayte tuff (LHT), which was most closely associated with the *A. boisei* remains, was dated to a less precise age of 1.468 Myr (weighted mean) with a standard error of 0.070 Myr, based on 14 plagioclase crystals. This latter date is consistent with the stratigraphic position of LHT between tuffs KRT and TBT. Details of the stratigraphy, lithologies and the ⁴⁰Ar/³⁹Ar data of these and other tuffs will be published elsewhere.

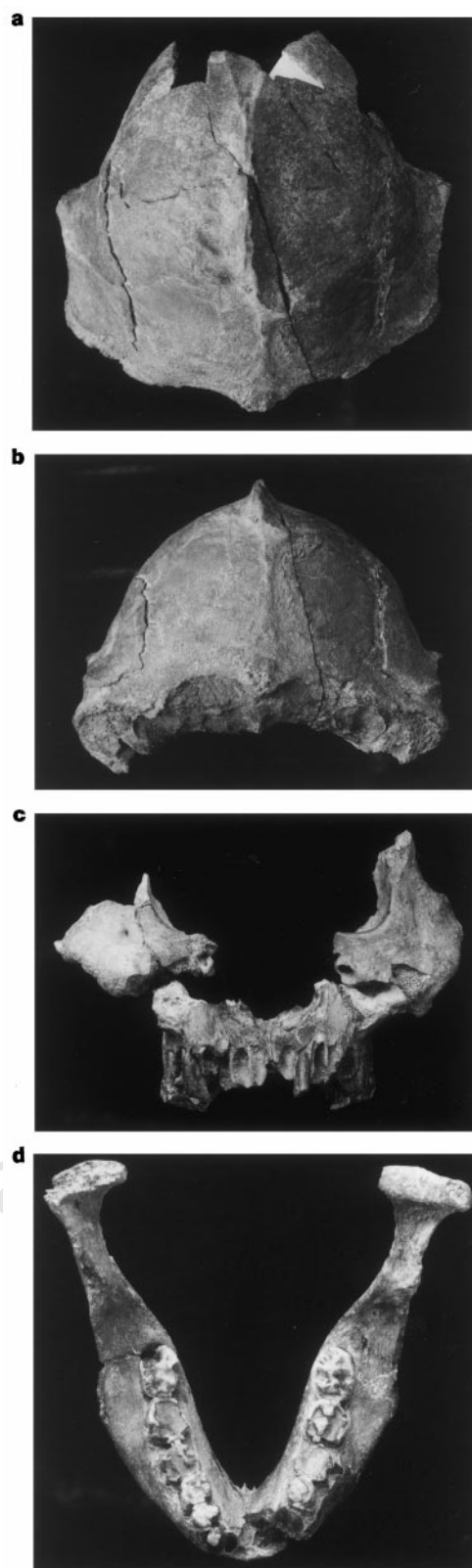


Figure 2 The KGA10-525 cranium and associated mandible. **a**, Superior view of cranial vault; **b**, posterior view of cranial vault; **c**, anterior view of face; **d**, superior view of mandible. All are approximately one-third natural size. The major craniomandibular measurements taken without reconstruction are as follows: bi-mastoid breadth, 151 mm; bi-porionic breadth, 143.5 mm; minimum malar height, 35 mm; maxillo-alveolar breadth at M2, 82 mm; internal alveolar breadth at M2, 47 mm; mandibular corpus height at M1, 41.5 mm (average of right and left); mandibular corpus breadth at M1, 29 mm (average of right and left).

Table 1 *Australopithecus boisei* remains recovered at Konso

Specimen number	Element	Discoverer	Year
KGA10-506	Left palate with dentition	A. Amzaye	1993
KGA10-525	Partial skull	A. Amzaye	1993
KGA10-565	Right upper M ¹	G. Suwa	1994
KGA10-570	Juvenile mandible	Y. Zeleke	1994
KGA10-900	Molar fragments	H. Nakaya	1994
KGA10-1455	Left parietal	K. Uzawa	1994
KGA10-1720	Left lower M ₃	B. Asfaw	1996
KGA10-2705	Right lower M ₂	Y. Haile-Selassie	1996
KGA10-2741	Molar fragments	K. Gelete	1996

Standard dental metrics are as follows: KGA10-506 M¹ mesiodistal length 14.2 mm (14.5 mm corrected for interproximal wear), buccolingual breadth 15.9 mm; KGA10-525 left lower M₂ buccolingual breadth 16.8 mm (estimate); left lower M₃ mesiodistal length 21.0 mm (21.2 mm corrected for interproximal wear), buccolingual breadth 17.7 mm; right upper M² buccolingual breadth 19.2 mm (estimate); right upper M³ buccolingual breadth 19.1 mm; KGA10-570 right lower M₁ mesiodistal length 16.2 mm, buccolingual breadth 14.9 mm; KGA10-1720 M₃ mesiodistal length 18.7 mm, buccolingual breadth 14.9 mm; KGA10-2705 M₂ buccolingual breadth 17.5 mm (estimate).

2,400 systematically collected non-hippopotamus mammalian specimens). In contrast, *A. boisei* remains are lacking from the more mesic Konso sites, despite a large combined sample of over 4,000 identifiable fossils. Because identical collecting procedures and criteria were used at all relevant localities, this distribution data at Konso is highly suggestive of actual association of *A. boisei* with dry grassland fauna. This calls for a re-evaluation of past interpretations of *A. boisei* habitat preference that were based on mixed locality and horizon tabulations⁷.

The main fossiliferous horizons of KGA10 lie stratigraphically between tuffs dated to 1.41 ± 0.02 and 1.43 ± 0.02 Myr (Fig. 1). *H. erectus* and early Acheulean artefacts occur in the same stratigraphic interval. Archeological occurrences from these levels exhibit an Acheulean industry characterized by a high frequency of picks. The abundance of Acheulean artefacts at Konso is in striking contrast to their absence in beds containing *H. erectus* and *A. boisei* east of Lake Turkana.

New *A. boisei* specimens include the KGA10-525 skull, the first known relatively complete associated cranium and mandible of the species (Table 1, Fig. 2). It comprises much of the facial skeleton and vault but lacks most of the frontal bone and anterior cranial base. Advanced cranial suture closure and dental wear indicate that the individual was in relatively old age. The KGA10-525 skull is large and presumably male. The bimastoid breadth of 151 mm is larger than that of six measurable *A. boisei* crania known from Olduvai and Turkana. A sagittal crest occurs at a mid-parietal position and a compound temporal/nuchal crest forms bilaterally. Cranial capacity is estimated to be approximately 545 cm³, slightly larger than hitherto recorded for *A. robustus* or *A. boisei*^{5,9}.

The face is superoinferiorly short as in KNM-ER 406, but more orthognathic as in O.H.5. The anteriorly situated zygomatic area is typical of robust *Australopithecus*. The mandible is at the large end of the known range of *A. boisei* variation⁴, and exhibits morphology common in that species. The worn dental arcade is considerably shorter in the maxilla than in the mandible, indicating negative overjet of the incisors. The premolars exhibit marked maxillary and mandibular wear discrepancy, with the maxillary counterparts showing excessive wear.

The overall morphology of the calvaria, facial skeleton and dentition of these Konso hominids is clearly that of robust *Australopithecus*, with additional features shared exclusively with *A. boisei* (Table 2). In particular, the modest dental collection recovered from KGA10 collectively shows the advanced occlusal morphology and dental proportions characteristic of that species¹⁰. Other features that strongly support the *A. boisei* taxonomic attribution include aspects of glenoid fossa and subnasal morphologies. We thus conclude that the KGA robust *Australopithecus* remains represent *A. boisei*. However, distinct morphologies are observed in the KGA10-525 skull, some of them hitherto

Table 2 Inter-taxon distribution of derived robust *Australopithecus* features and phenetic condition of Konso *A. boisei* remains*

	<i>A. boisei</i>	<i>A. aethiopicus</i>	<i>A. robustus</i>	Konso
1. Anterior emphasis of sagittal crest	+	—	+	(+)
2. Strong postorbital constriction	+	+	+	+
3. Extensive squamosoparietal suture overlap	+	(+)	?	+
4. Deep glenoid fossa	+	—	+	+
5. Temporomandibular joint lateral to cranial vault	+	+	(+)	(+)
6. Posteriorly oriented entoglenoid	+	—	—	+
7. Postglenoid merged with tympanic	+	—	+	—
8. Coronally oriented petrous	+	(+)	+	+
9. Vertically deep and plate-like tympanic	+	(+)	+	+
10. External auditory meatus obliquely elliptic	+	—	—	+
11. Laterally inflated mastoid with distinct supramastoid sulcus	+	+	+	+
12. Occipitomarginal sinus	+	(—)	+	—†
13. Anterior position of zygomatic maxillary plane	+	+	+	+
14. Anterior orientation of zygomatic frontal process	+	+	+	—
15. High position of zygomatic temporal process	+	+	+	+
16. Visor-like zygomatic morphology	+	—	—	—
17. Zygomatic prominence	—	(+)	+	+
18. Zygomaticomaxillary fossa	—	—	+	+
19. Zygomaticomaxillary step	—	(+)	+	—
20. Absence of canine fossa and maxillary fossula	+	+	—	+?
21. Anterior pillar	—	—	+	—?
22. Low position of infraorbital foramen	+	+	+	—
23. Thick palate	+	+	+	+
24. Recessed anterior nasal spine	+	(+)	+	+
25. Extensive overlap of palate roof and alveolar process	+	(+)	(+)	+
26. Deep palate	+	—	(+)	(+)
27. Median maxillary torus	—	—	+	(+)
28. Absolutely and relatively tall mandibular ramus	+	?	+	+
29. Relatively thick and inflated mandibular corpus	+	+	+	+
30. Wide extramolar sulcus	+	+	+	+
31. Mental protuberance	+	(+)	+	+
32. Anterior median symphyseal crest	+	—	—	+
33. Absolutely large postcanines	+	+	(+)	+
34. Relatively small anterior teeth	+	(+)	(+)	+
35. Specialized P ₃ morphology	+	+	+	+
36. Specialized P ₄ morphology	+	—	—	+
37. Molarized premolar roots	+	+	+	+
38. Lower molars with large distal cusps	+	(+)	(+)	+
39. Thick enamel	+	+	+	+
40. Flat molar wear	+	+	+	+

* These characters have been considered to be derived features characteristic of *A. boisei*, *A. robustus* and/or *A. aethiopicus*^{2-6,10,11,27-29}, relative to the *A. afarensis* and/or *A. africanus* conditions. Symbols + and — indicate presence and absence of the feature, respectively. Parentheses indicate intermediate or notably variable conditions. Because of arbitrary boundaries of presence/absence criteria, variability within species, and possible correlation among features, we caution against a numerical cladistic application of these tabulated data. Rather, this character list is meant to demonstrate the phenetic status of the KGA10-525 remains with respect to features suggested to have diagnostic value in robust *Australopithecus* taxonomy. † Endocranially there is no trace of either occipital or marginal sinuses, although impressions for the transverse sinus are also lacking. Nevertheless, a bilaterally developed sigmoid sinus groove of 8 mm diameter, twice the size of the same feature in O.H.5, suggests that venous drainage was by the transverse-sigmoid route.

unrecorded in *A. boisei* and/or other robust *Australopithecus* species.

The KGA10-525 sagittal crest commences behind bregma and continues past lambda. This posterior emphasis is known only in the *A. aethiopicus* cranium (KNM-WT 17000)^{3,11}, although the latter specimen exhibits an exaggerated crest projection not seen in the Konso skull. Approximately half of the temporomandibular joint is placed under the neurocranium, and the preglenoid plane is well developed medially. This may relate to the broad endocranial cavity of KGA10-525. The palate is broad and anteroposteriorly short, producing a *Homo*-like shape at an absolutely large size. The relatively laterally facing zygomatic surface is not the 'visor' morphology attributed to *A. boisei*⁶. Rather, delimited anterior and lateral zygomatic portions are more reminiscent of *A. robustus*. The preserved medial end of the infraorbital plane slopes posteromedially, suggesting a depressed superior circumnasal region.

The advanced age of the KGA10-525 individual, with extensive dental wear and unusual occlusion, may have influenced craniofacial remodelling. However, the wide and shallow palate is also seen in the younger KGA10-506 specimen, albeit to a lesser degree. Other features are unlikely to be correlative of this individual's advanced age. Crania with equivalent dental wear are known in *A. robustus*, but these do not show exaggerated zygomatic features. At least one *A. boisei* cranium, KNM-ER 13750, has a relatively advanced suture closure but exhibits the *A. boisei* 'visor'-like zygomatic. Finally, our survey of modern African apes, including individuals with mild negative overjet, revealed no obvious age-related morphological changes, apart from those evident in the alveolar and

subnasal regions. Analogy with modern African apes suggests that the strong medial infraorbital concavity of KGA10-525 is likely to be a reflection of individual variation in craniofacial hafting.

The Konso *A. boisei* specimens represent the best assemblage of this species convincingly dated between 1.4 Myr and 1.5 Myr. Ill-preserved cranial material from Chesowanja^{12,13} and the mandible from Lake Natron¹⁴ are also commonly considered to derive from the same time horizon, although these are not securely dated^{13,15}. The abundant Turkana and less-extensive Olduvai *A. boisei* fossils derive predominantly from the 1.5–2.0 Myr time period.^{4,5,16} The comparatively large cranial capacity of KGA10-525 follows an evolutionary trend suggested to have occurred in this lineage⁵. However, in most features commonly interpreted to be *A. boisei* masticatory specializations, the Konso sample exhibits a condition consistent with suggested stasis in that species¹⁶.

Polymorphism within *A. boisei* from the Turkana basin has recently been emphasized by Walker and his colleagues, who caution against a simplified typification of the species on the basis of the few relatively complete known specimens⁵. The Konso *A. boisei* skull provides further evidence of considerable intraspecific variation. Features that are outside, or at the end of, known Olduvai/Turkana *A. boisei* ranges of variation include zygomatic topography and palate depth. Other features that completely exceed known ranges of variation of both *A. boisei* and South African *A. robustus* include the strong lateral orientation of the zygomatic, the extreme superior circumnasal concavity, the broad palate and nasal aperture, the high position of the infraorbital foramen, and the form of the sagittal crest. In particular, lack of the visor-like facial

morphology contradicts Rak's dichotomization of East and South African robust *Australopithecus*⁶.

It is possible that the KGA *A. boisei* population as a whole was morphologically distinct from that of the Turkana basin, despite a geographical proximity of approximately 200 km. The Konso macromammalian fauna exhibits significant differences from the Turkana counterparts at both 1.4 Myr and 1.9 Myr levels. Dominant bovids and suids at Konso include *Kobus cf. sigmoidalis*, *Damaliscus niro*, *Parmularius angusticornis*, a short-horned *Syncerus*, *Tragelaphus strepsiceros*, *Kolpochoerus majus*, *Kolpochoerus limnetes/olduvaiensis*, *Notochoerus aff. euilus*, and *Metridiochoerus hopwoodi*. Most of these taxa are rare or absent at equivalent Turkana time horizons, or are different at the infraspecific level^{17,18}. Similarly, the abundant Acheulean assemblages found at KGA are lacking at East Turkana¹⁹. Thus, distinct differences between Konso and Turkana are seen in *A. boisei*, the fauna and archeological remains. Because all *Australopithecus* species hypodigms come from one dominant and a few subsidiary site samples, the extent of the polytypic nature of early hominid species may have been overlooked.

The importance of the Konso *A. boisei* fossils lies in their possession of uniquely derived features of that species, thus allowing secure taxonomic attribution at the species level, while otherwise extending the known range of *A. boisei* variation in many features. The striking deviations of the Konso morphology from the known Olduvai/Turkana *A. boisei* condition demonstrate that not all *A. boisei* features traditionally considered as part of a hyper-robust masticatory adaptation are of functional significance or decisive systematic value. This has important implications for early-hominid systematics and functional interpretations.

It is likely that many details of craniodental features used in evaluations of early hominids vary between populations in a manner consistent with random genetic drift. Excessive atomization of morphological features and their individual evaluations may then lead to erroneous phylogenetic and simplistic functional interpretations. We hypothesize that the facial morphology of the Konso skull, unique in the known *A. boisei* hypodigm, represents intraspecific variation, either individual, age-related, temporal and/or geographic. The striking distinctiveness of the Konso skull and the morphological trends common to both palates (KGA10-506 and KGA10-525) suggest that geographic variation is particularly influential. Geographically patterned, inter-basinal variation may be revealed further as more *A. boisei* specimens become available. Meanwhile, as some features once thought to characterize South African *A. robustus* are now shown to occur in East Africa, the case for robust *Australopithecus* monophyly is strengthened.

The taxonomic splitting of the ~2.0-Myr early-*Homo* hypodigm is a scheme rapidly gaining wide acceptance^{20–23}, and the newly proposed taxon *Australopithecus bahrelghazali* has been claimed to be distinct from *A. afarensis*²⁴. These interpretations rely heavily on specimen-specific morphological differences. Such differences have been taken as indicating cladogenesis, but the Konso discoveries combined with modern anthropoid variation^{25,26} suggest that considerable morphological polymorphism and/or polytypism may have been present among all early hominids. Splitting of taxa at the species level, emphasizing minor morphological differences, would then be unwarranted. We predict that, as new hominid-bearing sedimentary basins are discovered and existing site samples are augmented, a more accurate picture of variation will emerge to challenge overly split phylogenies of hominid species. □

Received 10 June; accepted 9 September 1997.

1. Leakey, L. S. B. A new fossil skull from Olduvai. *Nature* **184**, 491–493 (1959).
2. Tobias, P. V. T. *Olduvai Gorge Vol. 2, The Cranium and Maxillary Dentition of Australopithecus (Zinjanthropus) boisei* (Cambridge Univ. Press, 1967).
3. Leakey, R. E. F. & Walker, A. New *Australopithecus boisei* specimens from East and West Lake Turnaka, Kenya. *Am. J. Phys. Anthropol.* **76**, 1–24 (1988).
4. Wood, B. Koobi Fora Research Project Vol. 4 *Hominid Cranial Remains* (Clarendon, Oxford, 1991).
5. Brown, B., Walker, A., Ward, C. V. & Leakey, R. E. New *Australopithecus boisei* calvaria from East Lake Turkana, Kenya. *Am. J. Phys. Anthropol.* **91**, 137–159 (1993).
6. Rak, Y. *The Australopithecine Face* (Academic, New York, 1983).

7. Shipman, P. & Harris, J. M. in *Evolutionary History of the "Robust" Australopithecines* (ed. Grine, F.) 343–381 (de Gruyter, New York, 1988).
8. Asfaw, B. et al. The earliest Acheulean from Konso-Gardula. *Nature* **360**, 732–735 (1992).
9. Holloway, R. L. Human paleontological evidence relevant to language behaviour. *Hum. Neurobiol.* **2**, 105–114 (1983).
10. Suwa, G., White, T. D. & Howell, F. C. Mandibular postcanine dentition from the Shungura Formation, Ethiopia: Crown morphology, taxonomic allocations, and Plio-Pleistocene hominid evolution. *Am. J. Phys. Anthropol.* **101**, 247–282 (1996).
11. Walker, A., Leakey, R. E., Harris, J. M. & Brown, F. H. 2.5-Myr *Australopithecus boisei* from west of Lake Turkana. *Nature* **322**, 517–522 (1986).
12. Carney, J., Hill, A., Miller, J. & Walker, A. Late australopithecine from Baringo district, Kenya. *Nature* **230**, 509–514 (1971).
13. Gowlett, J. A. J., Harris, J. W. K., Walton, D. & Wood, B. A. Early archaeological sites, hominid remains and traces of fire from Chesowanja, Kenya. *Nature* **294**, 125–129 (1981).
14. Leakey, L. S. B. & Leakey, M. D. Recent discoveries of fossil hominids in Tanganyika at Olduvai and near Lake Natron. *Nature* **202**, 5–7 (1964).
15. Thouveny, N. & Taieb, M. Etude paleomagnetique des formations du Plio-Pleistocene de la Peninj (ouest du lac Natron, Tanzanie), limites de l'interpretation magnetostratigraphique. *Sci. Geol. Bull.* **40**, 57–70 (1987).
16. Wood, B., Wood, C. & Konigsberg, L. *Paranthropus boisei*: An example of evolutionary stasis? *Am. J. Phys. Anthropol.* **95**, 117–136 (1994).
17. Harris, J. M. *Koobi Fora Research Project Vol. 2, The Fossil Ungulates: Proboscidea, Perissodactyla, and Suidae* (Clarendon, Oxford, 1983).
18. Harris, J. M. *Koobi Fora Research Project Vol. 3, The Fossil Ungulates: Geology, Fossil Artiodactyls and Palaeoenvironments* (Clarendon, Oxford, 1991).
19. Leakey, M. G. & Leakey, R. E. *Koobi Fora Research Project Vol. 1, The Fossil Hominids and an Introduction to Their Context, 1968–1974* (Clarendon, Oxford, 1978).
20. Wood, B. Origin and evolution of the genus *Homo*. *Nature* **355**, 783–790 (1992).
21. Rightmire, G. P. Variation among early *Homo* crania from Olduvai Gorge and the Koobi Fora region. *Am. J. Phys. Anthropol.* **90**, 1–33 (1993).
22. Schrenk, F., Bromage, T. G., Betzler, C. G., Ring, U. & Juwayeyi, Y. Oldest *Homo* and Pliocene biogeography of the Malawi rift. *Nature* **365**, 833–836 (1993).
23. Kimbel, W. H., Johanson, D. C. & Rak, Y. Systematic assessment of a maxilla of *Homo* from Hadar, Ethiopia. *Am. J. Phys. Anthropol.* **103**, 235–262 (1997).
24. Brunet, M. et al. *Australopithecus bahrelghazali*, une nouvelle espèce d'Hominide ancien de la région de Koro Toro (Tchad). *C. R. Acad. Sci.* **322**, 907–913 (1996).
25. Albrecht, G. H. & Miller, J. in *Species, Species Concepts, and Primate Evolution* (eds Kimbel, W. H. & Martin, L. B.) 123–161 (Plenum, New York, 1993).
26. Uchida, A. *Peabody Museum Bulletin 4. Intra-species Variation among the Great Apes: Implications for Taxonomy of Fossil Hominoids* (Harvard Univ. Press, Cambridge, MA, 1996).
27. du Brul, E. L. Early Hominid feeding mechanisms. *Am. J. Phys. Anthropol.* **47**, 305–320 (1977).
28. Kimbel, W. H., White, T. D. & Johanson, D. C. in *Evolutionary History of the "Robust" Australopithecines* (ed. Grine, F.) 259–268 (de Gruyter, New York, 1988).
29. McCollum, M. A., Grine, F. E., Ward, S. C. & Kimbel, W. H. Subnasal morphological variation in extant hominoids and fossil hominids. *J. Hum. Evol.* **24**, 87–111 (1993).
30. WoldeGabriel, G. et al. Ecology and temporal placement of early Pliocene hominids at Aramis, Ethiopia. *Nature* **371**, 330–333 (1994).

Acknowledgements. The Konso Paleoanthropological Research Project has been conducted under permission from the Center for Research and Conservation of Cultural Heritage (CRCCH), Ministry of Information and Culture, Ethiopia. We thank the CRCCH and the National Museum of Ethiopia for allowing us to undertake our research; the Bureau of Culture and Information of the Southern Nations, Nationalities, and People's regional Government of Ethiopia for support; staff of the Bureau of Culture and Information and of CRCCH for participating in and contributing to the field research; A. Amzaye for discovering the KGA10-525 skull; the administration of the Konso Special Administration District and the Konso people for their support; A. Ademassu, T. Assebewerk, T. Becker, H. Gilbert, C. Guillemot, Y. Haile-Selassie, F. C. Howell, B. Latimer, S. Simpson, M. Umer and Y. Zeleke for field participation, laboratory assistance, and/or suggestions; P. Snow for access to the LANL Geology/Geochemistry Group microprobe facility and assistance; and R. Holloway for the cranial-capacity estimate of KGA10-525. This work was funded by the Mitsubishi Foundation and the Japan Ministry of Education, Science, Sports and Culture.

Correspondence and requests for materials should be addressed to G.S. (suwa@biol.s.u-tokyo.ac.jp).

Decrease of cardiac chaos in congestive heart failure

Chi-Sang Poon & Christopher K. Merrill

Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

The electrical properties of the mammalian heart undergo many complex transitions in normal and diseased states^{1–7}. It has been proposed that the normal heartbeat may display complex non-linear dynamics, including deterministic chaos^{8,9}, and that such cardiac chaos may be a useful physiological marker for the diagnosis^{10–12} and management^{13,14} of certain heart trouble. However, it is not clear whether the heartbeat series of healthy and diseased hearts are chaotic or stochastic^{15–17}, or whether cardiac chaos represents normal or abnormal behaviour¹⁸. Here we have used a highly sensitive technique, which is robust to random noise, to detect chaos¹⁹. We analysed the electrocardiograms from a group of healthy subjects and those with severe congestive heart failure (CHF), a clinical condition associated with a high risk of

sudden death. The short-term variations of beat-to-beat interval exhibited strongly and consistently chaotic behaviour in all healthy subjects, but were frequently interrupted by periods of seemingly non-chaotic fluctuations in patients with CHF. Chaotic dynamics in the CHF data, even when discernible, exhibited a high degree of random variability over time, suggesting a weaker form of chaos. These findings suggest that cardiac chaos is prevalent in healthy heart, and a decrease in such chaos may be indicative of CHF.

One of the main problems in studying the nonlinear dynamical behaviour of the heartbeat (and of any empirical time series in general) is the inevitable presence of random noise, which can often lead to false positive or negative identifications of chaos when traditional methods of nonlinear dynamics analysis are used. Although the effect of noise may be lessened to some extent by using long data records, this approach is limited by the possible non-stationarity of the heartbeat series^{20,21}.

To circumvent these difficulties, we have used a nonlinear systems identification technique¹⁹ to detect time-dependent and disease-dependent changes in the chaotic dynamics of the heartbeat. The technique identifies nonlinear determinism in a time series by iteratively generating a family of polynomial autoregressive models. The null hypothesis (namely, that the time series is stochastic with linear dynamics) is rejected if there is at least one nonlinear model that provides a significantly better fit to the data in a parsimonious manner than linear autoregressive models of all dynamical order. The statistical test is highly robust and sensitive, in that it is resistant to noise contamination and is applicable to short time series (with <1,000 points). This technique provides a highly specific test for deterministic chaos in that the null hypothesis is not readily rejected in the presence of random noise unless the underlying system is chaotic and/or has a fractal phase-space structure¹⁹. Indeed, simulation studies showed that the level of noise corruption that could be tolerated by such a chaotic test was directly related to the magnitude of the positive Lyapunov exponent, a measure of the degree of chaos in the underlying noise-free data (M. Barahona and

C.-S.P., unpublished observations). Moreover, the computational algorithm uses a recursive scheme for nonlinear systems identification and is highly efficient (typically, analysis of a 1,000-point time series takes just seconds on a desktop computer). The combination of specificity, sensitivity and computational efficiency of the chaotic test allowed a systematic statistical evaluation of the presence of deterministic chaos in multiple short segments of an extended heartbeat series, an analytical approach that would otherwise be impractical.

The data sets of interest were derived from an established database^{17,20,22} and consisted of continuous Holter records (at 128–250 samples s⁻¹) of heartbeat intervals from 8 healthy subjects and 11 CHF patients (age range, 22–71 years). At the time of the experiment, none of the subjects was on any medication that would affect the heart rate. To obviate possible non-physiological disturbances associated with wakefulness, data obtained only during nocturnal hours (totally 20,000 beats each) were used in the analysis. Subjects in both groups were selected on the basis of stability of the mean heart rate and limited number of ectopic beats and undetected beats. The heartbeat series were detrended using linear regression, but otherwise no preprocessing of the data was performed.

The heartbeat series of the CHF patients were characterized by the distinct presence of low-amplitude, low-frequency intermittent oscillations and decrease in temporal variability (Fig. 1), as reported previously^{22,23}. Generally, such an increase in regularity and decrease in temporal variability may reflect a decrease in either the noise or chaotic contents of the time series. To distinguish these possibilities, we first divided each heartbeat series into 40 contiguous short segments (of 500 beats each) to minimize the effect of non-stationarity of the data. We then applied the chaotic test to each data segment and evaluated the frequencies of linear and nonlinear episodes in each subject.

Nonlinear dynamics were manifested under the above test in both healthy and CHF heartbeat data (Fig. 2a, b). Because the chaotic test is conservative in the presence of noise (that is, predisposing to a

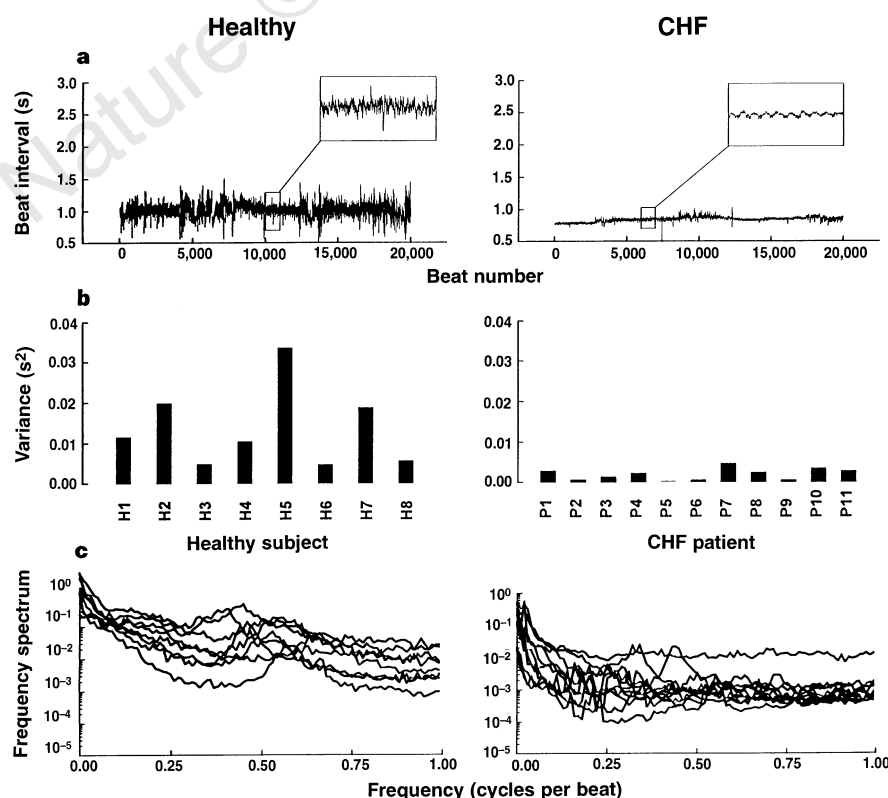


Figure 1 Characteristics of heartbeat data. **a**, Representative heartbeat series from a healthy subject (left) and a CHF patient (right). All heartbeat data were obtained during sleep and consisted of 20,000 beats. Insets show the self-similar character of the healthy heartbeat series and the decreased complexity and increased predictability of the CHF data in the form of low-amplitude, low-frequency intermittent oscillations. **b**, Variance of heartbeat intervals in all healthy subjects (left) and all CHF patients (right). Note the decreased variability of heartbeat in the CHF patient. **c**, Frequency spectrum of heartbeat intervals in all healthy subjects (left) and all CHF patients (right). Each frequency spectrum was obtained by taking a 2,000-point Fourier transform of the corresponding heartbeat series, averaged over all data segments, and then integrated over a 10-point bin in the frequency scale. The lack of sustained periodic or quasiperiodic components is indicated by the absence of frequency peaks in both subject groups.

linear model unless the data have a significant chaotic component), this finding suggests that the heartbeat series were intrinsically chaotic in both subject groups. The absence of other forms of nonlinear dynamics (such as limit cycles or invariant loops) is also indicated by the lack of constantly periodic or quasiperiodic components in the heartbeat series and corresponding frequency spectrum of all subjects (Fig. 1). Furthermore, the possible contribution of non-chaotic fractal attractors to the nonlinear dynamics may also be excluded, as such attractors occur only under certain special circumstances²⁴, which are unlikely in the cardiac system.

However, deterministic chaos was detected consistently in nearly

all data segments in the healthy group, but the detection rate was markedly decreased (by 20–80%) in all but one of the CHF patients (Fig. 2b). Because all data were from subjects during nocturnal hours, and the CHF data were even less variable than the normal data (Fig. 1), it is highly unlikely that the low detection rate in the CHF group was caused by increased noise corruption. To confirm this, we repeated the test with data segments of 2,000 beats each to increase the signal content. Again, all healthy heartbeat data with the longer segmentation proved strongly chaotic, but the corresponding CHF data still exhibited relatively low detection rates (Fig. 2c). Because a weakly chaotic time series is generally more susceptible to noise corruption than is a strongly chaotic one (one with a large positive Lyapunov exponent), the low detection rate in the CHF group suggests that the intensity of cardiac chaos was decreased in CHF patients.

Another way of measuring the changes in intensity of chaos in a noise-corrupted time series is to compare the variability of the resulting nonlinear models for those data segments that tested positive for deterministic chaos. Again, we make use of the property that the chaotic dynamics in a noise-corrupted time series is more readily identifiable if it has a higher signal-to-noise ratio. Thus the nonlinear coefficient estimates for the CHF data are expected to be less variable if the lower temporal variability of the CHF data (Fig. 1) reflected a decrease in noise level instead of a decrease in chaotic content.

To test this hypothesis, we computed the nonlinear coefficients for all chaotic segments of the healthy and CHF heartbeat series. Indeed, the estimated nonlinear coefficients were considerably more variable in the CHF patients than in healthy subjects (Fig. 3). This criterion again held in all but one of the CHF patients and for the group as a whole. The increased parameter variability in the CHF group, together with a corresponding decrease in detection rate for

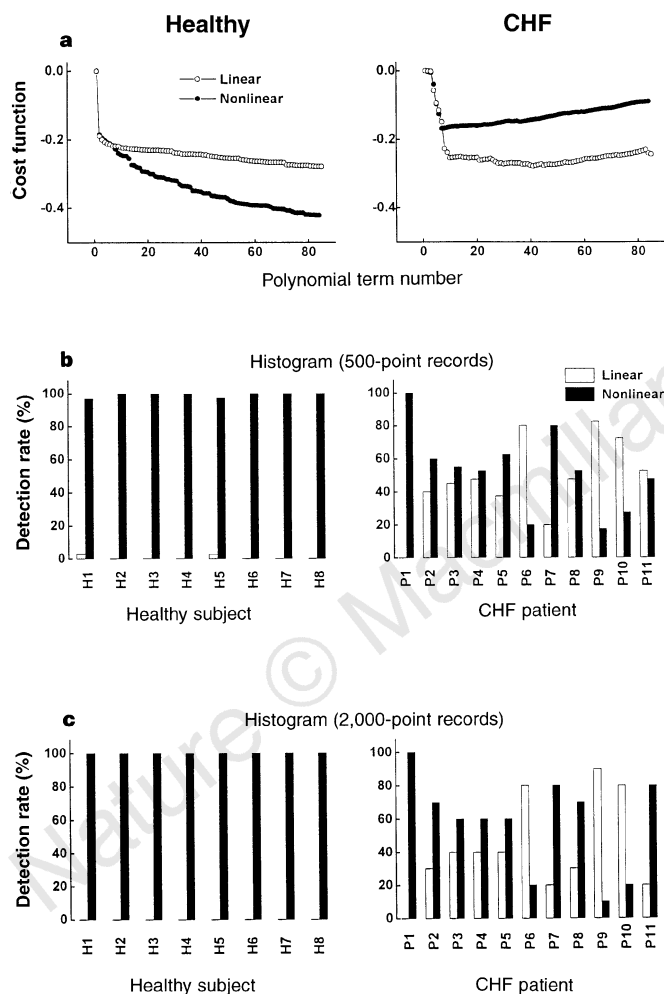


Figure 2 Chaos analysis of heartbeat series. The data were fitted with dynamical models of linear and polynomial forms with increasing order. Nonlinear dynamics was indicated if the best nonlinear model provides a better fit to the data than the best linear model with a similar number of polynomial terms. Model selection was based upon the *F*-ratio test for residuals at the 1% significance level. This technique is conservative in that it predisposes to a linear model (null hypothesis) in the presence of random noise unless the system is chaotic¹⁹. **a**, Examples of linear and nonlinear model fits of a 500-beat data segment for the same healthy subject (left) and CHF patient (right) shown in Fig. 1a. The cost function is of the form¹⁹: $C(r) = \log_e \epsilon(r) + r/N$, where r is the number of polynomial terms, $\epsilon(r)$ is the residual error, and N is the length of the time series. Both the cost function and the *F*-ratio test yielded a nonlinear model for the healthy subject and a linear model for the CHF subject. **b**, Histograms of linear and nonlinear model selection for all 500-beat data segments based on the above test in healthy subjects (left) and CHF patients (right). Note the high detection rates for chaos (nearly 100%) in the healthy group and the relatively low detection rates in CHF group (except patient P1). **c**, As **b**, but using 2,000-beat data segments.

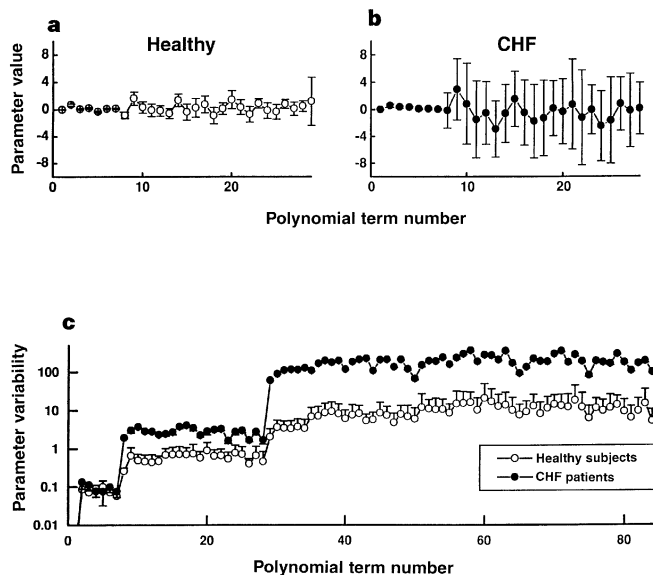


Figure 3 Variability of parameter estimates corresponding to all 500-beat data segments that were identified as chaotic in Fig. 2. Data shown are estimates of the leading nonlinear coefficients (mean $\pm$ s.d.) from the same healthy subject (**a**) and CHF patient (**b**) shown in Fig. 1a. Note the increased variability (s.d.) of estimated coefficients in the CHF patient. **c**, Average s.d. values (mean $\pm$ s.e.) of the first 80 nonlinear coefficient estimates for the healthy and CHF groups are plotted on a semi-logarithmic scale. The parameter estimates were significantly more variable in the CHF group than in the healthy group. All CHF patients except P1 exhibited similar increased parameter variability.

chaos (Fig. 2), confirm that the degree of cardiac chaos was decreased in the CHF patients.

These critical tests strongly support the suggested prevalence of cardiac chaos in healthy subjects^{8,9}. Moreover, our results indicate that cardiac chaos persists in CHF patients, albeit less strongly than in healthy subjects. The intermittent heartbeat oscillations characteristic of these patients²² (Fig. 1) suggest that they may be at the brink of intermittency, a common route to and out of chaos²⁵. Whereas the effect of noise contamination of the data precluded the reliable detection of chaos with previous approaches, we have used this property to evaluate statistically the changes in cardiac chaos with heart disease. Such a statistical approach has been made possible by the sensitivity, specificity and computational efficiency of the chaotic test¹⁹.

Our results do not reveal the mechanisms of cardiac chaos and its recession in heart failure; indeed abnormalities in left ventricular and autonomic system functions may all contribute to a decrease in complexity of the heartbeat nonlinear dynamics in CHF patients²². Nevertheless, of all the subjects tested, only one CHF patient failed both diagnostic criteria, corresponding to a type-I and type-II diagnostic error of 9% and 0%, respectively. Such remarkable consistency of the chaotic tests, together with the efficiency of the computational algorithm, suggest that such indices of chaos may be used as a specific, non-invasive and on-line diagnostic test for heart disease, and a possible indicator of imminent ventricular fibrillation²⁶. □

Received 20 May; accepted 1 August 1997.

- Smith, J. M. & Cohen, R. J. Simple finite-element model accounts for wide range of cardiac dysrhythmias. *Proc. Natl Acad. Sci. USA* **81**, 233–237 (1984).
- Chialvo, D. R. & Jalife, J. Nonlinear dynamics of cardiac excitation and impulse propagation. *Nature* **330**, 749–752 (1987).
- Chialvo, D. R., Gilmour, R. F. & Jalife, J. Low-dimensional chaos in cardiac tissue. *Nature* **343**, 653–657 (1990).
- Jalife, J. *Ann. NY Acad. Sci.* **591** (1990).
- Davidenko, J. M., Pertsov, R. S., Baxter, W. & Jalife, J. Stationary and drifting spiral waves of excitation in isolated cardiac muscle. *Nature* **355**, 349–351 (1989).
- Winfree, A. T. Electrical turbulence in three-dimensional heart muscle. *Science* **266**, 1003–1006 (1994).
- Glass, L. Dynamics of cardiac arrhythmias. *Phys. Today* **40–45** (1996).
- Goldberger, A. L. Is the normal heartbeat chaotic or homeostatic? *News Physiol. Sci.* **6**, 87–91 (1991).
- Sugihara, G., Allan, W., Sobel, D. & Allan, K. D. Nonlinear control of heart rate variability in human infants. *Proc. Natl Acad. Sci. USA* **93**, 2608–2613 (1996).
- Denton, T. A., Diamond, G. A., Helfant, R. H., Khan, S. & Karagueuzian, H. Fascinating rhythm: A primer on chaos theory and its application to cardiology. *Am. Heart J.* **120**, 1419–1440 (1990).
- Skinner, J. E., Goldberger, A. L., Mayer-Kress, G. & Ideker, R. E. Chaos in the heart: Implications for clinical cardiology. *Biotechnology* **8**, 1018–1024 (1990).
- Goldberger, A. L. Nonlinear dynamics for clinicians: Chaos theory, fractals and complexity at the bedside. *Lancet* **347**, 1312–1314 (1996).
- Garfinkel, A., Spano, M. L., Ditto, W. L. & Weiss, J. N. Controlling cardiac chaos. *Science* **257**, 1230–1235 (1992).
- Garfinkel, A., Weiss, J. N., Ditto, W. L. & Spano, M. L. Chaos control of cardiac arrhythmias. *Trends Cardiovasc. Med.* **5**, 76–80 (1995).
- Kaplan, D. T. & Cohen, R. J. Is fibrillation chaos? *Circ. Res.* **67**, 886–892 (1990).
- Kanters, J. K., Holstein-Rathlou, N.-H. & Agner, E. Lack of evidence for low-dimensional chaos in heart rate variability. *J. Cardiovasc. Electrophysiol.* **5**, 591–601 (1994).
- Turcott, R. G. & Teich, M. C. Fractal character of the electrocardiogram: Distinguishing heart-failure and normal patients. *Ann. Biomed. Eng.* **24**, 269–293 (1996).
- Glass, L. Is cardiac chaos normal or abnormal? *J. Cardiovasc. Electrophysiol.* **1**, 481–482 (1990).
- Barahona, M. & Poon, C.-S. Detection of nonlinear dynamics in short, noisy time series. *Nature* **381**, 215–217 (1996).
- Peng, C. K., Havlin, S., Stanley, H. E. & Goldberger, A. L. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos* **5**, 82–87 (1995).
- Ivanov, P. C. et al. Scaling behaviour of heartbeat intervals obtained by wavelet-based time-series analysis. *Nature* **383**, 323–327 (1996).
- Goldberger, A. L., Rigney, D. R., Mietus, J., Antman, E. W. & Greenwald, S. Nonlinear dynamics in sudden cardiac death syndrome: heart rate oscillations and bifurcations. *Experientia* **44**, 983–987 (1988).
- Casolo, G., Balli, E., Taddei, T., Amuhasi, J. & Gori, C. Decreased spontaneous heart rate variability on congestive heart failure. *Am. J. Cardiol.* **64**, 1162–1167 (1989).
- Grebogi, C., Ott, E., Pelikan, S. & Yorke, J. A. Strange attractors that are not chaotic. *Physica D* **13**, 261–268 (1984).
- Pomeau, Y. & Manneville, P. Intermittent transition to turbulence in dissipative dynamical systems. *Commun. Math. Phys.* **74**, 189–197 (1980).
- Skinner, J. E., Pratt, C.-M. & Vybiral, T. A. Reduction in the correlation dimension of heartbeat intervals precedes imminent ventricular fibrillation in human subjects. *Am. Heart J.* **125**, 731–743 (1993).

Acknowledgements. We thank A. L. Goldberger and R. G. Mark for discussions and comments on the manuscript, and A. L. Goldberger and J. E. Mietus for providing the heartbeat data. This work was supported by grants from the National Heart, Lung and Blood Institute, National Science Foundation, and Office of Naval Research.

Correspondence and requests for materials should be addressed to C.-S.P. (e-mail: cpoon@mit.edu).

A specific neural substrate for perceiving facial expressions of disgust

M. L. Phillips*, A. W. Young†, C. Senior*, M. Brammer‡, C. Andrew§, A. J. Calder†, E. T. Bullmore‡, D. I. Perrett||, D. Rowland||, S. C. R. Williams§, J. A. Gray† & A. S. David*

* Department of Psychological Medicine, King's College School of Medicine and Dentistry and Institute of Psychiatry, 103 Denmark Hill, London SE5 8AZ, UK

† Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, UK

§ Neuroimaging Unit, ‡ Brain Image Analysis Unit, ¶ Department of Psychology, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK

|| School of Psychology, University of St Andrews, Fife KY16 9JU, UK

Recognition of facial expressions is critical to our appreciation of the social and physical environment, with separate emotions having distinct facial expressions¹. Perception of fearful facial expressions has been extensively studied, appearing to depend upon the amygdala^{2–6}. Disgust—literally ‘bad taste’—is another important emotion, with a distinct evolutionary history⁷, and is conveyed by a characteristic facial expression^{8–10}. We have used functional magnetic resonance imaging (fMRI) to examine the neural substrate for perceiving disgust expressions. Normal volunteers were presented with faces showing mild or strong disgust or fear. Cerebral activation in response to these stimuli was contrasted with that for neutral faces. Results for fear generally confirmed previous positron emission tomography findings of amygdala involvement. Both strong and mild expressions of disgust activated anterior insular cortex but not the amygdala; strong disgust also activated structures linked to a limbic cortico–striatal–thalamic circuit. The anterior insula is known to be involved in responses to offensive tastes. The neural response to facial expressions of disgust in others is thus closely related to appraisal of distasteful stimuli.

We aimed to demonstrate distinct neural substrates for perception of two emotions, fear and disgust, replicating previous observations of a link between fear perception and amygdala activation^{5,6}, and examining the substrate for perception of disgust. It was postulated that perception of facial expressions of disgust would involve structures implicated in the appreciation of offensive stimuli. A cortico–striatal–thalamic circuit has been identified in primates¹¹, which may be involved in responses to emotive stimuli. There is clinical evidence for the probable involvement of some of these structures in appreciation of disgust: impaired recognition of disgust from facial expressions has been reported both in patients with symptomatic Huntington's disease¹², and presymptomatic carriers of the Huntington's gene¹³.

Subjects viewed grey-scale pictures of faces from a standard set¹⁴ depicting disgusted, fearful and neutral expressions. There were two levels of intensity (75 and 150%) for the facial expressions of disgust and fear for each individual face, and one level for the neutral expression, all produced by computer graphical manipulation of the prototype of each expression¹⁵ (Fig. 1). As a neutral face, we used an image with a slightly (25%) happy expression (see Methods). Subjects viewed blocks of emotional (disgusted or fearful) faces alternating with neutral faces in blocks. There were four separate experiments, in a randomized order, incorporating an alternating (neutral/emotional) design for each emotion (fear/disgust) and intensity of expression (mild, 75%; strong, 150%). After presentation of each face, subjects made a decision as to its sex by pressing one of two buttons with the right thumb. The sex decision task was chosen to allow an identical task and response across all conditions, and to permit comparison to a previous study of fear which also

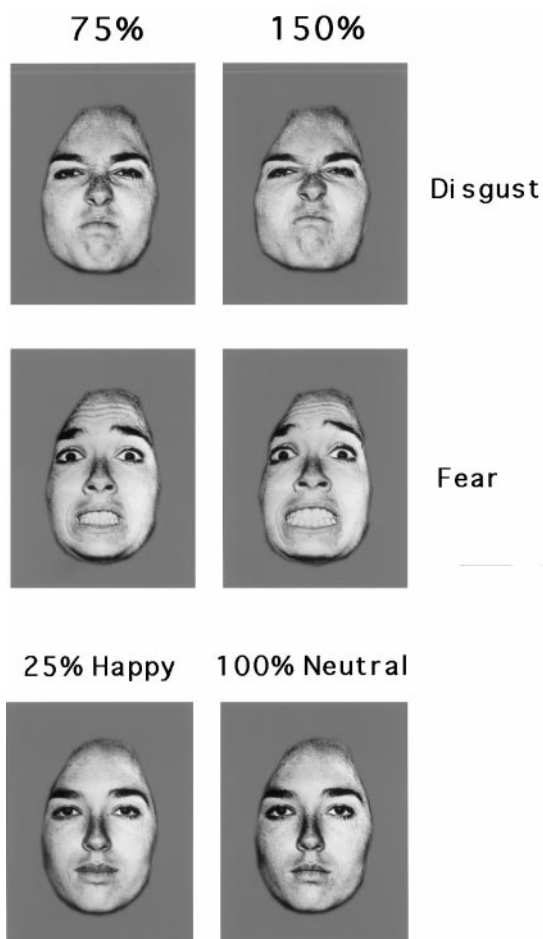


Figure 1 Faces from a standard set¹⁴ were computer-transformed¹⁵ to create two levels of intensity of expressed fear and disgust. Examples of faces depicting 100% neutral, 75 and 150% disgust, and 75 and 150% fear are demonstrated, together with an example of a stimulus depicting a mildly happy expression (75% neutral and 25% happy) which was used as the neutral baseline.

used this procedure⁶. We assumed (and our data confirm) that there would be a neural response to the face's displayed emotion even though subjects were not explicitly asked about it. Subjects were not informed that the aim of the study was to investigate responses to emotional expression.

Activation was demonstrated in the left amygdala for perception of fearful facial expression at the 75% intensity level with a voxel-wise probability of type I (false-positive) error $P < 0.004$. At this level of significance testing, we expect less than 6 false-positive voxels in a slice of the image comprising about 1,500 voxels in total. At the 150% level of fearful expression, we could not demonstrate activation of the amygdala at this relatively conservative level of significance testing. As the amygdala has been shown to be responsive to fear in other studies, we adopted a region of interest approach to analysis of these data. We tested the 18 voxels representing the amygdala complex bilaterally, with voxel-wise probability of type I error $P < 0.02$; at this level, we expect less than one false-positive voxel in the right and left amygdala regions. This approach demonstrated significant activation in the right amygdala for perception of 150% fearful facial expression (see Table 1a, b for details of generic brain activation foci in response to fearful facial expression). The lack of a dose-response effect for perception of fearful facial expression is discussed below.

The principal focus of interest in our study was the neural

Table 1 Main activated brain regions in the different experiments

BA/region	Side	Tal. x*	Tal. y*	Tal. z*	No activated voxels
(a) 75% Fearful faces versus neutral faces					
Insula	L	-40	-25	4	8†
Amygdala	L	-26	-14	-13	3†
		-26	-6	-7	1†
(b) 150% Fearful faces versus neutral faces					
Putamen	R	23	8	-7	7†
		29	11	20	5†
Amygdala	R	20	-11	-13	4‡
(c) 75% Disgusted faces versus neutral faces					
Anterior-mid Insula	R	35	-11	-2	5†
		38	-6	4	3†
32/Medial frontal cortex	R	17	-39	9	5†
(d) 150% Disgusted faces versus neutral faces					
32/Medial frontal cortex	L	-3	42	15	14†
18/Peristriae cortex	L	-14	-72	-7	11†
Anterior insula	R	35	31	9	10†
		46	11	9	8†
	L	-32	22	15	6†
23/Posterior cingulate gyrus	R	3	-53	15	8†
46/Dorsolateral prefrontal cortex	R	35	42	15	7†
19/Parastriae cortex	R	29	-75	15	5†
Thalamus	R	26	-31	15	5†
37/Inferior-posterior temporal cortex	L	-26	-50	-7	5†
		-40	-47	-7	4†
22/Superior temporal cortex	L	-43	-36	15	4†
21/Middle temporal cortex	R	46	-28	-2	4†
Putamen	R	23	3	9	4†
(e) Regions with significantly different activation for 150% versus 75% disgust					
Anterior insula	R	38	17	9	16†
18/Peristriae cortex	L	-20	-72	-7	8†

* Talairach co-ordinates refer to the voxel with the maximum FPQ (fundamental power quotient) in each regional cluster.

† Probability of false activation of each voxel in the generic brain map over all seven subjects was 0.004.

‡ Probability of false activation of each voxel in the amygdala region of interest over all seven subjects was 0.02.

response to facial expressions of disgust. The most striking finding for perception of facial expressions of disgust was activation in the right insula ($P < 0.004$), but not the amygdala (Table 1c, d and Fig. 2). Activation in the right anterior insula was significantly greater ($P < 0.004$) for perception of the 150% compared with the 75% disgust intensity (Table 1e and Fig. 3).

We have demonstrated a neural substrate for perception of facial expressions of disgust involving primarily the anterior insula, but, unlike fear, not the amygdala. The anterior insula is connected to the ventro-posterior-medial thalamic nucleus, and has been identified in primates as gustatory cortex¹⁶, containing neurons that respond to pleasant and unpleasant tastes¹⁷. In humans, anterior insula activation has been demonstrated while tasting salt in a functional imaging study¹⁸. Insula activation has also been demonstrated during perception of aversive stimuli such as pain¹⁹. Activation of the anterior insula during perception of facial expressions of disgust suggests that appreciation of visual stimuli depicting other's disgust is closely linked to the perception of unpleasant tastes and smells⁸. We also demonstrated activation in medial frontal cortex (Brodmann area 32, BA 32), anatomically connected with orbitofrontal cortex²⁰, during appreciation of both intensities of disgust, and in the right putamen and thalamus for the higher intensity of disgust. The emotional response to visceral, offensive stimuli may involve a limbic circuit connecting orbitofrontal cortex with ventral striatum

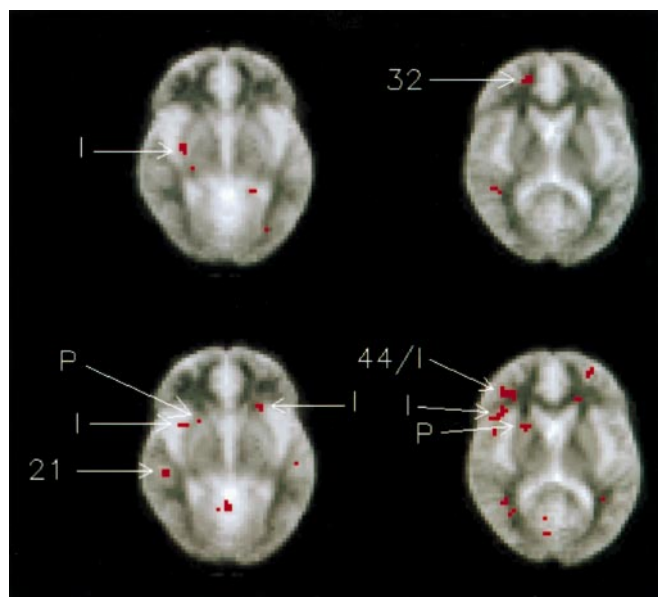


Figure 2 Generic brain activations in seven right-handed normal subjects during perception of faces depicting 75% (top row) and 150% (bottom row) disgust intensity. The grey-scale template was calculated by voxel-by-voxel averaging of the individual EPI images of all subjects, following transformation into Talairach space. The transverse sections in each experiment are at 2 mm below (left) and 9 mm above (right) the AC-PC line (right side of the brain on the left side of each section, and vice versa). Major regions of activation (probability of false activation <0.004) for perception of faces depicting 75% disgust versus a neutral expression are demonstrated in the right insula (I) and right medial frontal cortex (BA 32); those for faces depicting 150% disgust versus a neutral expression are demonstrated in the right and left anterior insula (I), right anterior insula bordering on inferior frontal cortex (BA 44), right putamen (P), and right middle temporal gyrus (BA 21).

and the mediodorsal thalamic nucleus¹¹. Our results suggest that structures linked with this circuit are involved in perception of facial expressions of disgust. Activation of posterior cingulate and visual cortex (Table 1d, e) may relate to arousal by an emotive visual stimulus²¹.

Unlike previously reported positron emission tomography (PET) findings⁶, there was no dose-response effect for amygdala activation with intensity of fearful facial expression. The time parameters differed between the PET study and our fMRI study; furthermore, a previous fMRI study has demonstrated rapid within-experiment habituation of amygdala response⁵, which could weaken the total extent of amygdala activation during the course of a five-minute experiment.

During the experimental procedure, subjects may have merely mimicked the facial expressions, rather than recognized the emotion depicted. As there was no activation demonstrated in the supplementary motor area, involved in movements of the mouth²², for any of the four experiments, this explanation for our results is unlikely.

The results demonstrate for the first time evidence for a differentiation between the neural responses to facial expressions of two negative emotions, fear and disgust. Disgust can be considered to be a response that evolved to lead to avoidance of contamination by offensive stimuli, and in particular, avoidance of ingestion of potentially harmful (for example, decayed) food⁸. We have demonstrated activation of the anterior insula, with its identified role as gustatory cortex, during appreciation of visual stimuli depicting expressions of disgust. Perception of others' disgust and that of taste

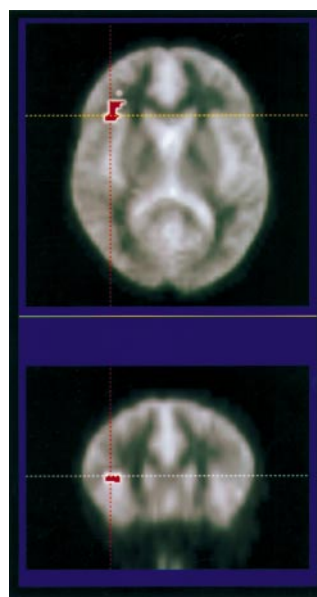


Figure 3 The difference image demonstrating significant ($P < 0.004$) differences in activation for perception of faces depicting 150% intensity of disgust (versus a neutral expression) and faces depicting 75% intensity of disgust (versus a neutral expression). The grey-scale template was as for Fig. 2. The largest region of activation was in the right anterior insula (Talairach coordinates 38, 17, 9), with twice the number of activated voxels compared with other regions of the difference image. Transverse ($z = 9$) and coronal ($y = 17$) sections are shown depicting this activation in the right insula.

appear, therefore, to have a similar neural substrate. It is interesting that the neural response to facial expressions of disgust in others should be linked to brain regions involved in gustatory responses. This suggests that our responses to others' disgust have, perhaps through associative learning between visual stimuli and taste²³, become closely linked to the appraisal of distasteful stimuli. □

Methods

Subjects. Seven right-handed healthy volunteers (five female and two male; mean age, 27 years; mean IQ estimate, 115) participated in the study. Exclusion criteria included history of brain injury and past and current psychiatric illness. No subject was taking regular medication.

Experimental design. Subjects participated in four 5-min experiments for presentation of emotional versus neutral facial stimuli. The faces of eight individuals (three male and five female) from a standard set of 'prototype' expressions of emotion¹⁴ were computer-transformed¹⁵ to create two levels of intensity of expressed fear and disgust (Fig. 1). At 75% intensity, the image was positioned with its features 75% along the continuum from neutral to the disgust prototype, or from neutral to the fear prototype. At 150% intensity, differences between the locations of facial features in the fear or disgust prototypes and a neutral expression were exaggerated by 50%. Studies of normal subjects have found that 150% images are recognized faster than 75% images for facial expressions of all basic emotions, showing the efficacy of this procedure in enhancing perceived emotion²⁴. In view of the fact that 100% neutral (muscles relaxed) faces from the standard set can appear slightly cold and threatening because it is conventional to signal approval in normal social interaction, we used as the neutral baseline stimulus a very slightly happy expression (75% neutral, 25% happy; Fig. 1). To familiarize subjects with the stimuli, they viewed all faces once, each presented for 3 s on a portable computer screen. In each experiment, the eight different faces depicting the same emotion and intensity of expression were presented one at a time on a computer screen in randomized order for 3 s each, followed by a 0.75-s interval in which the screen was blank. This was followed by presentation of eight neutral faces of the same eight individuals in a similar way. Stimuli were presented 3.5 m from the subject, subtending visual angles of ten degrees horizontally and eight degrees vertically. Each experiment comprised ten separate 30-s presentation phases, alternating between emotional (phase A) and neutral (phase B) stimuli, with the first presentation, either emotional (A) or neutral (B), counterbalanced across experiments for each subject, and also across subjects. The presentation order of the four experiments was also counterbalanced across subjects. Subjects made a decision as to the sex of each face by pressing one of two buttons with the right thumb. To allow for practice with the button box, subjects were presented with eight 100% neutral faces before each of the four different experiments. Accuracy of judgement of

sex for all faces by the seven subjects was near 100% (mean response over all subjects and experiments: 96.6%, range: 81.3–100%).

Image acquisition and analysis. Echoplanar MR brain images were acquired using a 1.5 Tesla GE Signa system (General Electric) retrofitted with advanced NMR hardware (ANMR) using a standard head coil. 100 T2*-weighted images depicting BOLD contrast²⁵ were acquired over 5 min (for each experiment) at each of 14 near-axial non-contiguous 5-mm-thick planes parallel to the intercommissural (AC-PC) line, providing whole-brain coverage: TE, 40 ms; TR, 3 s; in-plane resolution, 3 mm; interslice gap, 0.5 mm. An inversion recovery EPI dataset was also acquired at 43 near-axial 3-mm-thick planes parallel to the AC-PC line: TE, 80 ms; TI, 180 ms; TR, 16 s; in-plane resolution, 3 mm; number of signal averages, 8. The periodic change in T2*-weighted signal intensity at the (fundamental) experimentally determined frequency of alternation between A and B conditions was analysed by pseudogeneralized least-squares (PGLS) fit of a sinusoidal regression model to the movement-corrected²⁶ time series at each voxel, yielding parametric maps of the squared amplitude of the response at the stimulus frequency divided by its standard error—the fundamental power quotient, FPQ (ref. 27). Each observed time series was randomly permuted ten times, and FPQ estimated as above in each randomized time series, to generate 10 randomized parametric maps of FPQ for each subject in each anatomical plane. To construct generic brain activation maps, showing brain regions activated over a group of subjects, observed and randomized parametric maps of FPQ estimated in each individual were first transformed into the stereotactic space of Talairach and Tournoux and smoothed by a gaussian filter with full width at half maximum of 11 mm (ref. 28). The median observed value of FPQ was then computed at each voxel in standard space and its statistical significance tested by reference to the null distribution of median FPQ computed from the identically smoothed and spatially transformed randomized maps. For a one-tailed test of size p , the critical value was the $100*(1-p)$ th percentile of the randomization distribution²⁹. To identify voxels that demonstrated significant difference in standardized power of response to faces that expressed disgust with different intensities, the observed difference in median FPQ between these two experimental conditions was computed at each voxel. Subjects were then randomly reassigned to one of two equal-sized groups and the difference in median FPQ between randomized groups was computed at each voxel³⁰. This process was repeated 64 times and the results were pooled over voxels to generate a null distribution for difference in median FPQ. For a two-tailed test of size p , the critical values were the $100*(1-p/2)$ th and $100*(p/2)$ th percentiles of the randomization distribution.

Received 26 March; accepted 8 July 1997.

- Ekman, P. An argument for basic emotions. *Cog. Emot.* **6**, 169–200 (1992).
- Adolphs, R. et al. Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature* **372**, 669–672 (1994).
- Adolphs, R. et al. Fear and the human amygdala. *J. Neurosci.* **16**, 7678–7687 (1995).
- Calder, A. J. et al. Facial emotion recognition after bilateral amygdala damage: differentially severe impairment of fear. *Cogn. Neuropsychol.* **13**, 699–740 (1996).
- Breiter, H. C. et al. Response and habituation of the human amygdala during visual processing of facial expression. *Neuron* **17**, 875–887 (1996).
- Morris, J. et al. A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature* **383**, 812–815 (1996).
- Darwin, C. *The Expression of the Emotions in Man and Animals* (University of Chicago Press, Chicago, 1965).
- Rozin, P. & Fallon, A. E. A perspective on disgust. *Psychol. Rev.* **94**, 23–41 (1987).
- Rozin, P., Lowery, L. & Ebert, R. J. Varieties of disgust faces and the structure of disgust. *Pers. Soc. Psychol.* **66**, 870–881 (1994).
- Vrana, S. R. The psychophysiology of disgust: differentiating negative emotional context with facial EMG. *Psychophysiology* **30**, 279–286 (1993).
- Alexander, G. E. et al. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* **9**, 357–381 (1986).
- Spengler, R. et al. Perception of faces and emotions: loss of disgust in Huntington's disease. *Brain* **119**, 1647–1665 (1996).
- Gray, J. M. et al. Impaired recognition of disgust in Huntington's disease gene carriers. *Brain* (in the press).
- Ekman, P. & Friesen, W. V. *Pictures of Facial Affect* (Consulting Psychologists, Palo Alto, 1976).
- Perrett, D. I., May, K. A. & Yoshikawa, S. Facial shape and judgements of female attractiveness. *Nature* **368**, 239–242 (1994).
- Rolls, E. T. in *Handbook of Clinical Olfaction and Gustation* (ed. Doty, R. L.) (Dekker, New York, 1994).
- Yaxley, S. et al. The responsiveness of neurons in the insular gustatory cortex of the macaque monkey is independent of hunger. *Physiol. Behav.* **42**, 223–229 (1988).
- Kinomura, S. et al. Functional anatomy of taste perception in the human brain studied with positron emission tomography. *Brain Res.* **659**, 263–266 (1994).
- Casey, K. L. et al. Positron emission tomographic analysis of cerebral structures activated specifically by repetitive noxious heat stimuli. *J. Neurophysiol.* **71**, 802–807 (1994).
- Barbas, H. Organization of cortical afferent input to orbitofrontal areas in the Rhesus monkey. *Neuroscience* **56**, 841–864 (1993).
- Kosslyn, S. M. et al. Neural effects of visualizing and perceiving aversive stimuli: a PET investigation. *NeuroReport* **7**, 1569–1576 (1996).

- Gentilucci, M. et al. Functional organization of inferior area 6 in the macaque monkey: I. Somatotopy and the control of proximal movements. *Exp. Brain Res.* **71**, 475–490 (1988).
- Rolls, E. T. et al. Orbitofrontal cortex neurons: Role in olfactory and visual association learning. *J. Neurophysiol.* **75**, 1970–1981 (1996).
- Calder, A. J., Young, A. W., Rowland, D. & Perrett, D. I. Computer-enhanced emotion in facial expressions. *Proc. R. Soc. B.* (in the press).
- Ogawa, S., Lee, T. M., Kay, A. R. & Tank, D. W. Brain magnetic resonance imaging with contrast dependent blood oxygenation. *Proc. Natl Acad. Sci. USA* **87**, 8868–8872 (1990).
- Friston, K. J. et al. Movement-related effects in fMRI time series. *Mag. Res. Med.* **35**, 346–355 (1996).
- Bullmore, E. T. et al. Statistical methods of estimation and inference for functional MR image analysis. *Mag. Res. Med.* **35**, 261–277 (1996).
- Talairach, J. & Tournoux, P. *Co-planar Stereotactic Atlas of the Human Brain* (Thieme, Stuttgart, 1988).
- Brammer, M. J. et al. Generic brain activation mapping in fMRI: a nonparametric approach. *Mag. Res. Imag.* (in the press).
- Edgington, E. S. *Randomisation Tests* (Dekker, New York, 1980).

Acknowledgements. M.L.P. is supported by an MRC clinical training fellowship; E.T.B. is supported by the Wellcome Trust. We thank H. Critchley for advice and for comments on the manuscript, and A. Simmons for technical support.

Correspondence and requests for materials to M.L.P. (e-mail: spmanlp@iop.bpmf.ac.uk).

Transmissions to mice indicate that 'new variant' CJD is caused by the BSE agent

M. E. Bruce*, R. G. Will†, J. W. Ironside†, I. McConnell*, D. Drummond*, A. Suttie*, L. McCardle†, A. Chree*, J. Hope‡, C. Birkett‡, S. Cousens§, H. Fraser* & C. J. Bostock‡

* Institute for Animal Health, BBSRC/MRC Neuropathogenesis Unit, West Mains Road, Edinburgh EH9 3JF, UK

† National CJD Surveillance Unit, Western General Hospital, Edinburgh EH4 2XU, UK

‡ Institute for Animal Health, Compton, Newbury, Berkshire RG20 7NN, UK

§ Department of Epidemiology and Population Science, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK

There are many strains of the agents that cause transmissible spongiform encephalopathies (TSEs) or 'prion' diseases. These strains are distinguishable by their disease characteristics in experimentally infected animals, in particular the incubation periods and neuropathology they produce in panels of inbred mouse strains^{1–4}. We have shown that the strain of agent from cattle affected by bovine spongiform encephalopathy (BSE) produces a characteristic pattern of disease in mice that is retained after experimental passage through a variety of intermediate species^{5–7}. This BSE 'signature' has also been identified in transmissions to mice of TSEs of domestic cats and two exotic species of ruminant^{6,8}, providing the first direct evidence for the accidental spread of a TSE between species. Twenty cases of a clinically and pathologically atypical form of Creutzfeldt–Jakob disease (CJD), referred to as 'new variant' CJD (vCJD)⁹, have been recognized in unusually young people in the United Kingdom, and a further case has been reported in France¹⁰. This has raised serious concerns that BSE may have spread to humans, putatively by dietary exposure. Here we report the interim results of transmissions of sporadic CJD and vCJD to mice. Our data provide strong evidence that the same agent strain is involved in both BSE and vCJD.

Transmissions to mice were set up from six typical sporadic cases of CJD (spCJD) and three cases of vCJD. All were homozygous for methionine at codon 129 of the 'prion protein' (PrP) gene, and none carried PrP gene mutations associated with familial disease. The spCJD cases included two dairy farmers (aged 61 and 64 years) who had had BSE in their herds and had therefore been potentially exposed to BSE-infected cattle or contaminated animal feed¹¹; two 'contemporary' cases (aged 55 and 57 years) with no known occupational exposure to BSE; and two 'historical' cases (aged 57 and 82 years) who had died in 1981 and 1983, before the onset of the

BSE outbreak. All of these spCJD cases were characterized by widespread spongiform vacuolation in the brain with few or no amyloid plaques. The vCJD cases (aged 29, 30 and 31 years) had clinical and neuropathological characteristics that were atypical for CJD⁹. The main distinguishing neuropathological features in these and other vCJD cases are an extensive deposition of PrP amyloid in the brain as large 'florid' plaques and a prominent involvement of the cerebellum.

Panels of three inbred mouse strains (RIII, C57BL and VM) and one cross (C57BL × VM) were challenged with CJD brain homogenates. Previous transmissions, using the same protocol, of BSE from eight unrelated cattle (Fig. 1b) and TSEs from three domestic cats (Fig. 1c), a greater kudu and a nyala (two exotic ruminants) have given a remarkably uniform pattern of incubation periods in these mice^{5,6,8}. The shortest incubation periods were seen in RIII mice, with means ranging from 302 to 335 days for transmissions from frozen brain samples. These isolates also produced strikingly similar patterns of vacuolar degeneration in the brains of infected mice, as represented by the 'lesion profile'^{5,6} (Fig. 2b, c). The lesion profile is a well-established semiquantitative method of measuring the targeting of vacuolation to different brain regions, and reliably discriminates between TSE strains in mice². In addition, the disease characteristics in mice injected with brain from two sheep, a goat and a pig that had been experimentally infected with BSE were very similar to those seen in direct BSE transmissions from cattle^{6,7}.

The BSE 'signature', based on both incubation periods and pathology, has only ever been seen in transmissions from animals

suspected or known to have been infected with BSE. It has never been seen throughout an extensive series of transmissions, set up in Edinburgh between 1963 and 1994, of other naturally occurring TSEs (35 sheep and two goats with scrapie, two mink with transmissible mink encephalopathy, and a mule deer with chronic wasting disease). For example, the incubation periods and lesion profiles seen in transmissions from six sheep with scrapie, collected since 1985, are shown in Figs 1a and 2a. Within the same timescale a further two sources of sheep scrapie failed to transmit to mice. In general, natural scrapie transmissions in our own laboratory and elsewhere have given variable results, probably reflecting variation in agent strain amongst the sheep sources^{12,13}.

At the time of writing, the transmissions of vCJD to mice have been in progress for 360 days. The RIII mice injected with all three vCJD sources have developed a progressive clinical disease very similar to BSE, with incubation periods in individual mice ranging from 288 to 351 days. The first signs were nervousness and hypersensitivity, followed by lethargy, weight loss, urinary incontinence and postural abnormalities. Excluding early intercurrent deaths, all RIII mice injected with vCJD have developed disease, with mean incubation periods up to a standard clinical endpoint of 304 ± 4 , 306 ± 6 and 310 ± 4 days ($\pm$ s.e.m.) for three sources (Fig. 1d), within but at the lower end of the range previously seen for BSE and related isolates⁶ (Fig. 1b, c). Clinical signs are now apparent in some of the C57BL mice, an observation that is also consistent with the BSE pattern (see Fig. 1b). Diagnosis was confirmed for all clinically affected mice by the presence of vacuolar degeneration in

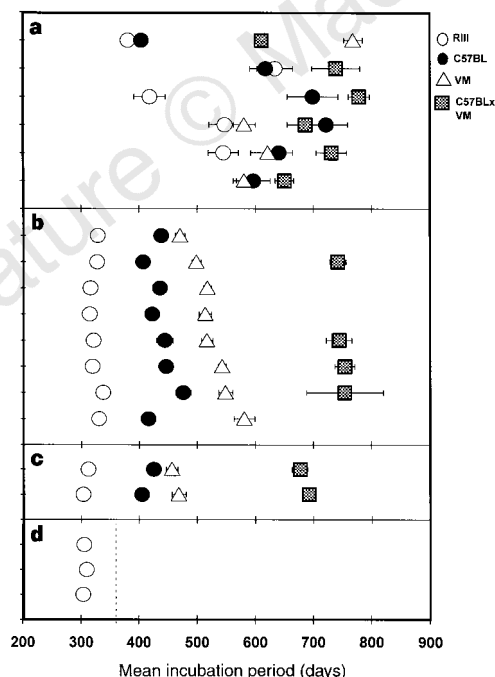


Figure 1 Incubation times in mice with spongiform encephalopathies. Incubation periods in RIII, C57BL, VM and C57BL × VM mice in transmissions of: **a**, natural scrapie from six sheep; **b**, BSE from eight cattle; **c**, FSE from two cats; and **d**, three cases of vCJD. C57BL × VM mice were not included in the first, third and fourth BSE transmission; missing symbols elsewhere indicate that no clinical disease was seen in these groups up to the natural lifespan of the mice. The vertical dotted line in **d** shows the current time after challenge in experiments still in progress. Data are mean $\pm$ s.e.m.

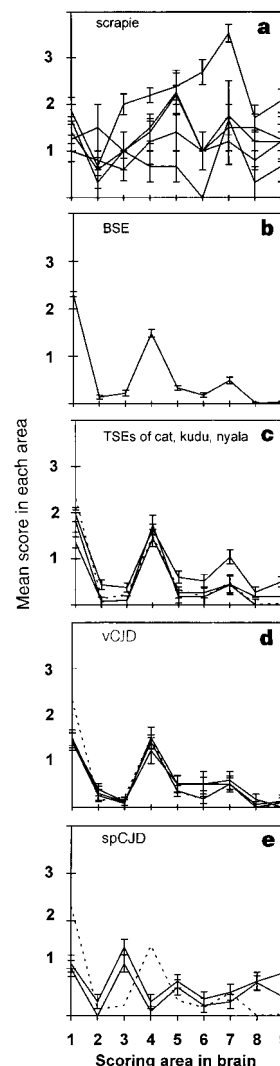


Figure 2 Lesion profiles for mice following transmission of spongiform encephalopathies. Lesion profiles are for RIII mice in transmissions of: **a**, natural scrapie from five of the six sheep from Fig. 1a ($n = 3-16$ mice per group; no clinical disease was seen in this mouse strain in the sixth transmission); **b**, BSE from the first four cattle ($n = 123$, pooled data); **c**, FSE from two cats ($n = 36$, pooled data) and TSEs from a greater kudu and a nyala ($n = 12$ and 11); **d**, vCJD from three sources ($n = 10$, 12 and 16) and **e**, spCJD from two sources, a farmer and a contemporary case ($n = 8$ and 9). The pooled BSE profile is shown as a dotted line in **c-e**. As white-matter vacuolation was not a prominent feature in any of these transmissions, only the grey-matter lesion profiles are shown. Vacuolation was scored on a scale of 0-5 in the following scoring areas: 1, dorsal medulla; 2, cerebellar cortex; 3, superior colliculus; 4, hypothalamus; 5, thalamus; 6, hippocampus; 7, septum; 8, retrosplenial and adjacent motor cortex; and 9, cingulate and adjacent motor cortex. Data are mean $\pm$ s.e.m.

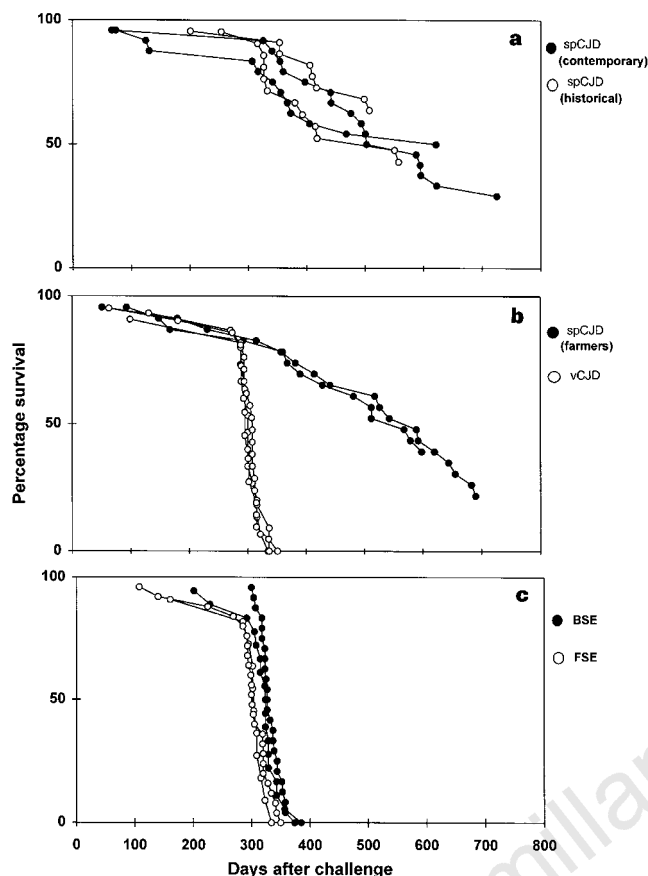


Figure 3 Survival curves for mice following transmission of spongiform encephalopathies. Survival curves are for female RIII mice in transmissions of: **a**, spCJD from cases with no known occupational exposure to BSE; **b**, spCJD from two farmers and vCJD from three sources; and **c**, BSE from the first two cattle sources and FSE from two cats. No distinction is made between mice dying with clinical signs of TSE infection and mice dying with intercurrent disease. Deaths up to 50 days after challenge, most of which were related to injection trauma, are excluded from the analysis.

the brain and for selected mice in all three experiments by the demonstration of relatively protease-resistant isoforms of PrP (PrP^{res}) in western blots of brain extracts and pathological accumulations of PrP in immunostained brain sections.

The neuropathology in clinically affected RIII mice with vCJD was also similar to that seen in RIII mice with BSE, consisting of a mild-to-moderate grey-matter vacuolation of the hypothalamus, medulla oblongata and septum, and a more severe vacuolation of the cochlear nucleus. Amyloid plaques were not a prominent feature of this pathology. The lesion profiles in RIII mice for the three sources (Fig. 2d) were very similar to each other and also to those in transmissions to RIII mice of BSE (Fig. 2b), TSEs of cats and exotic ruminants (Fig. 2c), and experimental sheep, goat and pig BSE^{6,7}, but differed markedly from those seen in transmissions from sheep with natural scrapie (Fig. 2a). Although results are so far only available for the RIII mouse strain, the striking similarity between vCJD and BSE in these mice, in terms of both incubation periods and pathology, is in itself strong evidence that the same strain of agent is involved in vCJD and BSE.

In contrast to the results with the vCJD sources, no clinical signs of neurological disease have yet been seen in any mice in the six transmissions of spCJD, although they have been in progress for between 600 and 800 days. Figure 3 shows survival curves for RIII

mice in these experiments, compared with survival curves in BSE, feline spongiform encephalopathy (FSE) and vCJD transmissions. No significant differences in median survival times were found between RIII groups challenged with the six spCJD sources, or between these groups and saline-injected controls. However, this does not indicate a failure to transmit spCJD, as vacuolar degeneration typical of TSE infection was seen in the brains of some mice dying with intercurrent disease in all six experiments, from about 400 days after challenge. This pathology has so far been seen in 130 of the 156 mice surviving beyond 500 days after injection for which brain was available for histopathological scrutiny. No such changes have been seen in the control mice of any age in this set of experiments, or in mice of the same strains injected with human brain homogenates from patients with amyotrophic lateral sclerosis or laryngeal carcinoma in a previously study¹⁴. Western blot and immunohistochemical analyses have demonstrated the accumulation of PrP^{res} in selected brains showing vacuolar pathology, confirming successful transmission of a TSE from all six spCJD sources.

A full analysis of the pathology in recipient mice in spCJD transmissions will be presented when these experiments are complete, but already several points can be made. Vacuolar degeneration has been seen in all four mouse strains. This pathology differs in severity between individual mice, but shows a consistent pattern between mouse strains and between spCJD sources. The earliest pathology is seen consistently in the superior colliculus and olfactory tract. In brains showing more widespread vacuolation there is also prominent involvement of the cerebral cortex, thalamus, hypothalamus, caudate nucleus and optic tract, a distribution unlike that seen in BSE transmissions to mice. As an illustration, Fig. 2e shows the lesion profiles for RIII mice killed with intercurrent disease or culled between 500 and 750 days after challenge with two spCJD sources (a farmer and a contemporary case). Although these profiles are not based on animals at the clinical endpoint of the disease, they clearly show a similarity in lesion distribution between the two sources of spCJD, and a difference between these sources and vCJD or BSE, particularly in scoring area 3, the superior colliculus. The results of these transmissions therefore provide no evidence of a link between CJD in dairy farmers and BSE.

A series of transmissions to mice of spCJD and familial human TSEs associated with mutations in the PrP gene have been reported in Japan¹⁵. Although different mouse strains were used in the Japanese series, the results for transmissions of spCJD from 129-methionine sources were broadly similar to ours in that transmission was achieved from all sources and mean incubation periods in recipient mice were long (573–863 days)¹⁵. Transmissions of the familial TSE Gerstmann–Straussler–Scheinker syndrome (GSS) were achieved from only one-third of the sources tested, but the mean incubation periods in successful transmissions were relatively short (237–517 days)¹⁵. Although some of these incubation periods were quite close to our results for vCJD in RIII mice, the pathology in mice with GSS was strikingly different as it included a prominent vacuolation of white-matter tracts¹⁶. The Japanese workers also reported the transmission of another human familial TSE, fatal familial insomnia (FFI), with a mean incubation period of 455 days in recipient mice¹⁷. The pathology in mice with FFI was indistinguishable from that in mice with spCJD in the Japanese series, apart from there being a more pronounced involvement of the thalamus.

Our results highlight several fundamental features of the TSEs previously established using experimental isolates^{3,4}. The consistency in transmission properties shows that the agent must interact with genetic factors in the host to control the timing and neuropathology of the disease with extraordinary precision. Different strains of agent (spCJD, vCJD) can be isolated from hosts with the same PrP amino-acid sequence (in this case, patients with the 129-methionine genotype) but, conversely, the same strain of agent can

be detected in hosts with different PrP sequences (so far the BSE 'signature' has been seen in transmissions from eight different species⁶). This clearly indicates that TSE agents carry some form of information that specifies strain-specific properties, but the molecular basis of this information is still a matter for speculation⁴.

It has been reported that vCJD can be distinguished from spCJD by the relative prominence of differently glycosylated forms of PrP^{res} and the molecular size of the unglycosylated form¹⁸. Samples from dairy farmers with CJD have given glycoform ratios resembling those from other cases of spCJD¹⁹. A similarity in glycoform patterns between vCJD and BSE has been presented as evidence of a link between the two¹⁸. However, a 'BSE-like' glycoform pattern has also been seen for experimental scrapie isolates that are unrelated to BSE²⁰ and for FFI in humans²¹. Therefore, although the analysis of PrP diversity provides a useful supplement to strain typing in mice, it is premature to draw conclusions concerning causative links between TSEs in different species on the basis of glycoform-ratio analysis alone. A full analysis of glycoform patterns in the present series of transmissions will be reported in due course.

In conclusion, strain typing based on transmission to mice has shown: that vCJD is caused by the same strain of agent that has caused BSE, FSE and TSEs in exotic ruminants; that vCJD is distinguishable from spCJD; and that CJD in two dairy farmers is of the spCJD type and is not linked to the causative agent of BSE. Epidemiological surveillance continues to indicate that vCJD is a new condition occurring almost exclusively in the UK. Our transmission studies, in combination with the surveillance data, provide compelling evidence of a link between BSE and vCJD. □

Methods

CJD inocula. The CJD challenge experiments were the first to be undertaken within a new category 3 containment facility at the Neuropathogenesis Unit in Edinburgh, in an environment in which no TSE-infected materials had been handled previously. New dedicated glassware and instruments were autoclaved at 136 °C for 1 h before use. Brain samples for transmission were collected, as far as possible, from areas showing maximum pathology, and stored at -20 °C. Samples were homogenized at 10% (w/v) concentration in sterile physiological saline and stored at -20 °C. Before homogenization of each sample, sterile physiological saline was run through the homogenizer and other glassware and frozen for later inoculation of the appropriate control group. For injection, thawed homogenates were resuspended by being drawn repeatedly through a series of graded needles.

CJD transmissions. Three inbred mouse strains and one cross were challenged: C57BL and RIII (both of the *Sinc*⁸⁷ or *Prn-p*^a genotype), VM (*Sinc*^{p7} or *Prn-p*^b genotype), and the F₁ cross between C57BL and VM^{22,23}. Groups of approximately 20 mice of each strain were injected by a combination of the intracerebral (20 µl) and intraperitoneal (100 µl) routes under halothane anaesthesia. For each transmission, six mice of each strain were injected with the appropriate saline sample by the same routes. Groups of uninjected control mice were also included. Mice were coded, examined daily throughout their lifespan, and formally scored for signs of neurological disease from 250 days after injection. Mice showing definite signs for two consecutive weeks were killed and incubation periods calculated as the interval between injection and this standard clinical endpoint²². All other mice were maintained to full lifespan, apart from small numbers culled at 700–750 days post-injection, to avoid loss of pathological material.

Histopathological and protein analysis. At post-mortem, a lateral third of each mouse brain was dissected aseptically and frozen at -20 °C for protein analysis and further passage. The remaining two-thirds of each brain was immersion fixed in 10% formol saline for 4 days, treated with 98–100% formic acid for 1 h to inactivate infectivity, and fixed in formol saline for a further 2 days. The brains were trimmed at standard coronal levels and paraffin embedded. Haematoxylin and eosin-stained sections 6 µm thick were prepared, randomly mixed with others from BSE, FSE, sheep scrapie and mouse-passaged scrapie transmissions, and coded for pathological assessment. Vacuolar changes were scored in nine grey-matter and three white-matter areas of brain for the construction of lesion profiles, as described²⁴. PrP in brain

sections was immunostained using a polyclonal antibody to mouse PrP, 1A8 (ref. 25), according to a published protocol²⁶. SDS–polyacrylamide gel electrophoresis and western blot analysis of brain tissue were used to confirm the presence of PrP^{res} (ref. 18).

Animal TSE transmissions. The CJD transmissions were compared with transmissions of TSEs from cattle, sheep, domestic cats, a greater kudu and a nyala; the results of some of these animal TSE transmissions have been included in previous publications^{5,6,8}. The design of these experiments was identical to that in the CJD transmissions, except that the source material from one cat, the greater kudu and the nyala was formol-fixed brain. Because transmissions from fixed tissues have resulted in prolonged incubation periods⁵, probably owing to loss of titre, the incubation period and survival data from these three transmissions are not included in Figs 1 and 3. However, as the pathology in experimentally infected mice is unaffected by the dose of TSE challenge, lesion profile data from the kudu and nyala transmissions are included in Fig. 2.

Statistical analyses. Statistical analyses were performed using the software package Stata, version 5.0. Kaplan–Meier survival curves were plotted and differences in survival between mice inoculated with material from different sources were compared using the log-rank test²⁷. Principal components analysis²⁸ was used to calculate summary measures of lesion profiles and to examine graphically the 'closeness' of the lesion profiles from different transmissions. This statistical analysis was in complete agreement with the subjective judgement of 'closeness' described in the text and will be documented in detail in a future publication.

Received 8 August; accepted 9 September 1997.

- Dickinson, A. G. & Meikle, V. M. H. Host-genotype and agent effects in scrapie incubation: change in allelic interaction with different strains of agent. *Mol. Gen. Genet.* **112**, 73–79 (1971).
- Fraser, H. & Dickinson, A. G. Scrapie in mice: agent-strain differences in the distribution and intensity of grey matter vacuolation. *J. Comp. Pathol.* **83**, 29–40 (1973).
- Bruce, M. E., McConnell, I., Fraser, H. & Dickinson, A. G. The disease characteristics of different strains of scrapie in *Sinc* congenic mouse lines: implications for the nature of the agent and host control of pathogenesis. *J. Gen. Virol.* **72**, 595–603 (1991).
- Bruce, M. E. Scrapie strain variation and mutation. *Br. Med. Bull.* **49**, 822–838 (1993).
- Fraser, H., Bruce, M. E., Chree, A., McConnell, I. & Wells, G. A. Transmission of bovine spongiform encephalopathy and scrapie to mice. *J. Gen. Virol.* **73**, 1891–1897 (1992).
- Bruce, M. *et al.* Transmission of bovine spongiform encephalopathy and scrapie to mice: strain variation and the species barrier. *Phil. Trans. R. Soc. Lond. B* **343**, 405–411 (1994).
- Foster, J. D., Bruce, M., McConnell, I., Chree, A. & Fraser, H. Detection of BSE infectivity in brain and spleen of experimentally infected sheep. *Vet. Rec.* **138**, 546–548 (1996).
- Fraser, H. *et al.* Transmission of feline spongiform encephalopathy to mice. *Vet. Rec.* **134**, 449 (1994).
- Will, R. G. *et al.* A new variant of Creutzfeldt–Jakob disease in the UK. *Lancet* **347**, 921–925 (1996).
- Chazot, G. *et al.* New variant of Creutzfeldt–Jakob disease in a 26-year-old French man. *Lancet* **347**, 1181 (1996).
- Cousens, S. N. *et al.* Sporadic Creutzfeldt–Jakob disease in the United Kingdom: epidemiological data from 1970–1996. *Br. Med. J.* **315**, 389–396 (1997).
- Dickinson, A. G. in *Slow Virus Diseases of Animals and Man* (ed. Kimberlin, R. H.) 209–241 (North-Holland, Amsterdam, 1976).
- Carp, R. I. & Callahan, S. M. Variation in the characteristics of 10 mouse-passaged scrapie lines derived from five scrapie-positive sheep. *J. Gen. Virol.* **72**, 293–298 (1991).
- Fraser, H., Behan, W., Chree, A., Crossland, G. & Behan, P. Mouse inoculation studies reveal no transmissible agent in amyotrophic lateral sclerosis. *Brain Pathol.* **6**, 89–99 (1996).
- Tateishi, J. Transmission of human prion diseases of rodents. *Semin. Virol.* **7**, 175–180 (1996).
- Tateishi, J., Ohta, M., Koga, M., Sato, Y. & Kuroiwa, Y. Transmission of chronic spongiform encephalopathy with kuru plaques from humans to small rodents. *Ann. Neurol.* **5**, 581–584 (1979).
- Tateishi, J. *et al.* First experimental transmission of fatal familial insomnia. *Nature* **376**, 434–435 (1995).
- Collinge, J., Sidle, K. C. L., Meads, J., Ironside, J. & Hill, A. F. Molecular analysis of prion strain variation and the aetiology of 'new variant' CJD. *Nature* **383**, 685–690 (1996).
- Hill, A. F., Will, R. G., Ironside, J. & Collinge, J. Type of prion protein in UK farmers with Creutzfeldt–Jakob disease. *Lancet* **350**, 188 (1997).
- Somerville, R. A. *et al.* Biochemical typing of scrapie strains. *Nature* **386**, 564 (1997).
- Telling, G. *et al.* Evidence for the conformation of the pathological isoform of the prion protein enciphering and propagating prion diversity. *Science* **274**, 2079–2082 (1996).
- Dickinson, A. G., Meikle, V. M. H. & Fraser, H. Identification of a gene which controls the incubation period of some strains of scrapie agent in mice. *J. Comp. Pathol.* **78**, 293–299 (1968).
- Westaway, D. *et al.* Distinct prion proteins in short and long scrapie incubation period mice. *Cell* **51**, 651–662 (1987).
- Fraser, H. & Dickinson, A. G. The sequential development of the brain lesions of scrapie in three strains of mice. *J. Comp. Pathol.* **78**, 301–311 (1968).
- Farquhar, C. F. *et al.* in *Transmissible Spongiform Encephalopathies* (eds Bradley, R. & Marchant, B.) 301–313 (Commission of the European Communities, Brussels, 1994).
- Bell, J. E. *et al.* Prion protein immunocytochemistry—UK five centre consensus report. *Neuropathol. Appl. Neurobiol.* **23**, 26–35 (1997).
- Kirkwood, B. R. *Essentials of Medical Statistics* (Blackwell, Oxford, 1988).
- Mardia, K. V., Kent, J. T. & Bibby, J. M. *Multivariate Analysis* (Academic, London, 1979).

Acknowledgements. We acknowledge the contribution of A. Dickinson, who pioneered TSE strain discrimination in the 1960s. We also thank M. Brady, F. Purdie, L. Hunter, K. Lamza, S. Mack and other staff at the Neuropathogenesis Unit for technical support. This work was supported by the Department of Health, MRC, BBSRC and MAFF.

Correspondence and requests for materials should be addressed to M.E.B. (e-mail: moira.bruce@bbsrc.ac.uk).

AMPA receptor-mediated regulation of a G_i-protein in cortical neurons

Yizheng Wang, Daniel L. Small, Danica B. Stanimirovic, Paul Morley & Jon P. Durkin

Cellular Neurobiology Group, Institute for Biological Sciences, National Research Council of Canada, Ottawa, Ontario K1A 0R6, Canada

Excitatory synaptic transmission in the central nervous system is mediated primarily by the release of glutamate from presynaptic terminals onto postsynaptic channels gated by *N*-methyl-D-aspartate (NMDA) and α -amino-3-hydroxy-5-methylisoxazole-4-propionate (AMPA) receptors^{1,2}. The myriad intracellular responses arising from the activation of the NMDA and AMPA receptors have previously been attributed to the flow of Ca²⁺ and/or Na⁺ through these ion channels¹⁻⁶. Here we report that the binding of the agonist AMPA to its receptor can generate intracellular signals that are independent of Ca²⁺ and Na⁺ in rat cortical neurons. In the absence of intracellular Ca²⁺ and Na⁺, AMPA, but not NMDA, brought about changes in a guanine-nucleotide-binding protein (G_i) that inhibited pertussis toxin-mediated ADP-ribosylation of the protein in an *in vitro* assay. This effect was observed in intact neurons treated with AMPA as well as in isolated membranes exposed to AMPA, and was also found in MIN6 cells, which express functional AMPA receptors but have no metabotropic glutamate receptors. AMPA also inhibited forskolin-stimulated activity of adenylate cyclase in neurons, demonstrating that G_i proteins were activated. Moreover, both G $\beta\gamma$ blockage and co-precipitation experiments demonstrated that the modulation of the G_i protein arose from the association of

G_i with the glutamate receptor-1 (GluR1) subunit. These results suggest that, as well as acting as an ion channel, the AMPA receptor can exhibit metabotropic activity.

Treatment of cortical neurons with AMPA, but not NMDA, caused a profound reduction in the pertussis toxin (PTX)-mediated ADP-ribosylation of a membrane protein of relative molecular mass 41,000 (*M_r* 41K) in a subsequent *in vitro* assay (Fig. 1a). This AMPA-induced effect was not observed in corresponding cytosolic fractions subjected to the same PTX assay (Fig. 1b). Western blot analysis using an antibody specific to G_i confirmed that the 41K protein was G_i, and that the reduction in ADP-ribosylation seen following AMPA exposure was therefore not due to the loss of membrane G_i (Fig. 1a). A similar reduction in the ADP-ribosylation of G_i was seen when neurons were exposed to AMPA in the absence of extracellular Ca²⁺ and Na⁺ (Fig. 1c). This AMPA-mediated effect on G_i was blocked by the selective AMPA-receptor antagonist CNQX (Fig. 1d), was not mimicked by the metabotropic glutamate receptor (mGluR) agonists *trans*-(1S,3R)-ACPD (Fig. 1c) or AP-4 (data not shown), and was mimicked by thrombin (Fig. 1c), an agent known to activate a G_i protein⁷. These results indicate that activation of an AMPA receptor modulates a G_i protein in cortical neurons. Because this AMPA-induced effect occurred in an environment which precluded the flow of Ca²⁺ and Na⁺ through the receptor, it provided the initial evidence that the reduced ADP-ribosylation of G_i mediated by AMPA was independent of the receptor's ion-channel activity.

Membranes isolated from unstimulated neurons were examined for the ability of AMPA to affect both the PTX-mediated ADP-ribosylation of membrane G_i and the forskolin-induced increase of cyclic AMP levels when added *in vitro*. Both AMPA (Fig. 2a) and thrombin (Fig. 2b), but not *trans*-(1S,3R)-ACPD or NMDA (data not shown), substantially reduced the ADP-ribosylation of G_i relative to untreated membranes. This inhibition by AMPA was concentration dependent with a threshold between 10–25 μ M AMPA (Fig. 2c), was not due to AMPA affecting membrane levels

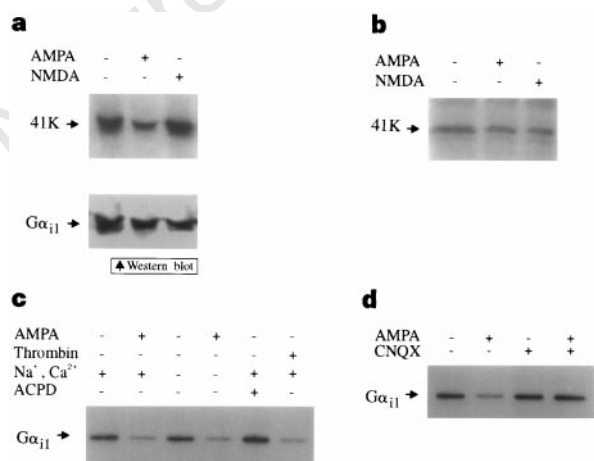


Figure 1 AMPA, added to intact neurons, inhibited the PTX-mediated ADP-ribosylation of G_i. Neurons were challenged for 3 min with the following agents, and the ability of PTX to ADP-ribosylate G_i in isolated membranes or cytosol was determined. **a, b**, Effects of 100 μ M AMPA or NMDA on ADP-ribosylation of membrane (**a**) and cytosolic (**b**) G_i. **c**, Effects of 1 unit thrombin or 100 μ M *trans*-(1S,3R)-ACPD on membrane G_i, and the influence of extracellular Ca²⁺ and Na⁺ on the effect of 100 μ M AMPA on G_i. **d**, The blockade of the effect of AMPA on membrane G_i by CNQX (100 μ M). Data are representative of three separate determinations.

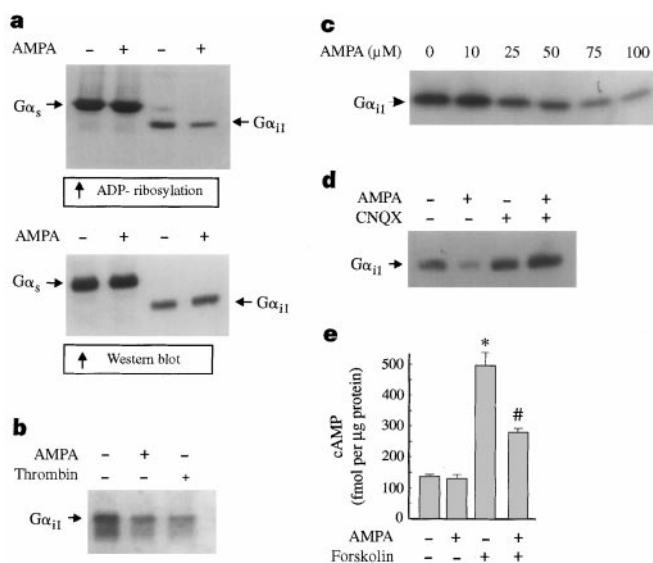


Figure 2 AMPA, added to membranes isolated from unstimulated neurons, reduced the PTX-mediated ADP-ribosylation of G_i and forskolin-stimulated adenylate cyclase activity. **a**, Effects of AMPA on G_i or G_s. **b**, Effects of AMPA or 1 unit thrombin on G_i. **c**, Concentration-dependent effects of AMPA on G_i. **d**, The reversal of the effect of AMPA on G_i by 100 μ M CNQX added 3 min before AMPA. Data are representative of three experiments. **e**, Inhibition of forskolin-stimulated adenylate cyclase activity by AMPA. The results of cAMP levels are mean $\pm$ s.e.m. of four independent experiments in duplicate. **P* < 0.0001 versus control; #*P* < 0.001 versus forskolin.

of $G\alpha_{i1}$ (Fig. 2a), and was observed even though the assay was performed in a buffer free of added Ca^{2+} and Na^{+} and containing 2 mM EGTA. In contrast, AMPA had no effect on the ability of cholera toxin (CTX) to ADP-ribosylate $G\alpha_s$ (Fig. 2a). The AMPA effect was blocked by CNQX (Fig. 2d), but not by the metabotropic receptor blocker MCPG (500 μ M) or the selective NMDA-channel blocker MK-801 (data not shown). AMPA had no effect on PTX, NAD^{+} or any other aspect of the ADP-ribosylation assay, as PTX-mediated ADP-ribosylation of a commercially available purified G_i protein was unaffected in the presence of a wide range of AMPA concentrations (data not shown). AMPA significantly inhibited forskolin-stimulated adenylate cyclase activity as measured by the reduction of cAMP levels (Fig. 3e).

The simplest explanation consistent with these results is that the binding of AMPA to its receptor activated the G_i protein and caused the dissociation of its α -subunit from the $\alpha\beta\gamma$ heterotrimeric complex, an event known to render the α -subunit less susceptible to PTX-mediated ADP-ribosylation^{8–10}. Two additional pieces of evidence indicate that this is indeed the case. First, the addition of exogenous $G\beta\gamma$ subunit to the *in vitro* ADP-ribosylation assay prevented the ability of AMPA to inhibit the ADP-ribosylation of $G\alpha_{i1}$ in isolated cortical membrane preparations (Fig. 3a). This would be expected because the free $G\alpha_{i1}$ generated through the action of AMPA reformed the heterotrimeric complex in the presence of excess $\beta\gamma$ subunits. Furthermore, co-precipitation studies demonstrated that AMPA triggered a rapid association between $G\alpha_{i1}$ and the AMPA receptor itself, as the anti- $G\alpha_{i1}$ antibody was able to detect the presence of $G\alpha_{i1}$ in immunocomplexes precipitated by an anti-GluR1 antibody from AMPA-challenged neurons, but not from control cultures (Fig. 3b). This AMPA-induced co-precipitation of $G\alpha_{i1}$ with GluR1 was independent of extracellular Ca^{2+} and Na^{+} (Fig. 3b), was blocked by CNQX (Fig. 3b), and was not found in cells exposed to either *trans*-(1S,3R)-ACPD or thrombin (Fig. 3c). These results indicated that the binding of AMPA to its receptor initiated an association of $G\alpha_{i1}$

with the GluR1 subunit, leading to the subsequent inhibition of ADP-ribosylation of $G\alpha_{i1}$ catalysed by PTX, and that this association was independent of the flow of Ca^{2+} and Na^{+} through the AMPA receptor.

It has been shown that AMPA can act as a weak agonist to mGluRs, specifically mGluR6 (ref. 11). It was therefore important to test the possibility that G_i activation by AMPA arose not from direct AMPA-receptor stimulation, but through the AMPA-induced activation of mGluRs, which would release $G\alpha_{i1}$ capable of associating with the AMPA receptor. However, several lines of evidence indicated that the effects of AMPA were not due to AMPA-mediated activation of mGluRs. First, *trans*-(1S,3R)-ACPD, an agonist for mGluRs, was unable to inhibit ADP ribosylation of $G\alpha_{i1}$ in intact neurons (Fig. 1c) or to cause the association of $G\alpha_{i1}$ with GluR1 (Fig. 3c) even though, at the same concentrations, it induced increases in internal Ca^{2+} concentration and reductions in cAMP levels in the neuronal cultures (data not shown). Moreover, the presence of mGluR6 was not detected in our cortical cultures using reverse transcription–polymerase chain reaction (RT–PCR) and western blot analyses (data not shown), consistent with reports that mGluR6 expression is restricted to rod bipolar cells of the retina^{12,13}. Second, thrombin, which inhibited ADP-ribosylation of $G\alpha_{i1}$ in the neurons (Fig. 1c), did not cause the association of $G\alpha_{i1}$ with GluR1 (Fig. 3c), indicating that GluR1 was not a general target for free $G\alpha_{i1}$. Third, AMPA, added to either isolated membranes (Fig. 2e) or intact neurons (data not shown) reduced forskolin-induced cAMP levels through the inhibition of adenylate cyclase activity, demonstrating that AMPA did indeed activate the G_i proteins. Finally, AMPA was also able to reduce the PTX-mediated ADP-ribosylation of $G\alpha_{i1}$ in MIN6 insulin-secreting pancreatic β -cells, a cell line possessing functional AMPA receptors^{14,15} but without mGluR1–7 subtypes, based on western blot and RT–PCR analyses (Fig. 4a, b). As was found with the neuronal membranes, AMPA, but not NMDA, inhibited PTX-mediated ADP-ribosylation of $G\alpha_{i1}$ in membranes prepared from unstimulated MIN6 cells

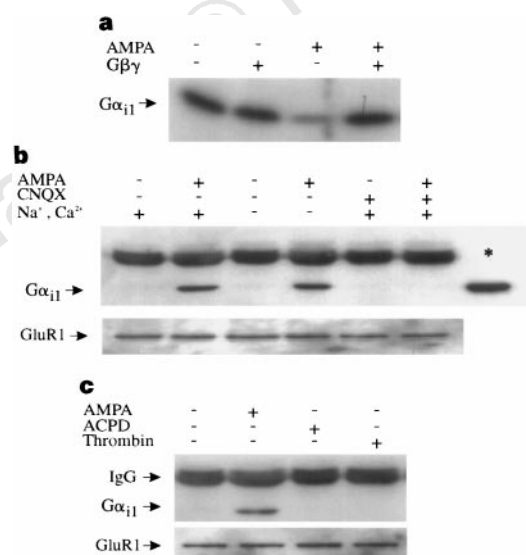


Figure 3 Reversion of AMPA-inhibited ADP-ribosylation of $G\alpha_{i1}$ by exogenous $G\beta\gamma$ subunit in neuronal membranes, and the detection of $G\alpha_{i1}$ in immunocomplexes precipitated by an anti-GluR1 antibody in neurons exposed to AMPA. **a**, Effects of exogenous $G\beta\gamma$ subunit on the AMPA-mediated inhibition of ADP-ribosylation of $G\alpha_{i1}$. **b**, The ability of AMPA to induce the association of $G\alpha_{i1}$ with the GluR1 subunit of the AMPA receptor. Asterisk indicates purified bovine brain G proteins (Calbiochem). **c**, Effects of AMPA, *trans*-(1S,3R)-ACPD or thrombin on the co-precipitation of $G\alpha_{i1}$ with GluR1.

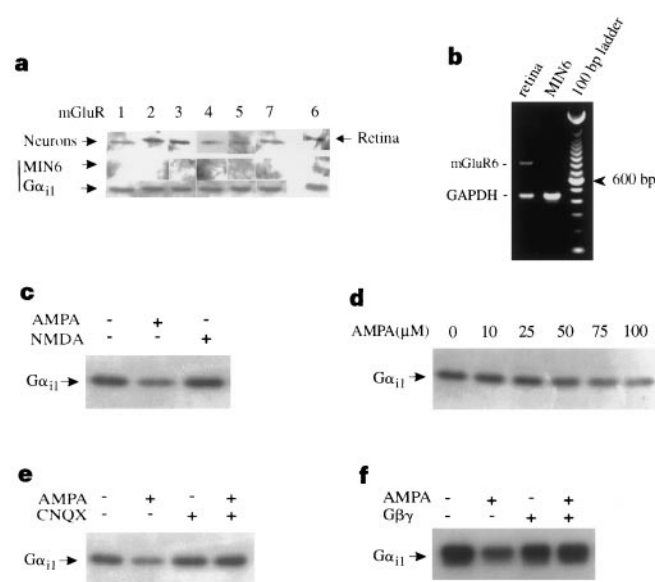


Figure 4 AMPA added to membranes isolated from unstimulated MIN6 cells, inhibited PTX-mediated ADP-ribosylation of $G\alpha_{i1}$. **a**, **b**, The absence of mGluRs in MIN6 cells as detected by western blot analysis (**a**) and RT–PCR (**b**). The PCR products were 806 bp (mGluR6) and 441 bp (GAPDH). **c**, Effects of AMPA or NMDA on $G\alpha_{i1}$. **d**, Concentration-dependent effects of AMPA on ADP-ribosylation of $G\alpha_{i1}$. **e**, Blockage of the AMPA effect on $G\alpha_{i1}$ by 100 μ M CNQX added 3 min before AMPA. **f**, Restoration of the inhibitory effect of AMPA on $G\alpha_{i1}$ by exogenous $G\beta\gamma$ subunit.

(Fig. 4c). This AMPA effect was concentration dependent (Fig. 4d), was inhibited by CNQX (Fig. 4e), and was reversed by the addition of exogenous G $\beta\gamma$ to the *in vitro* ADP-ribosylation assay (Fig. 4f). Thus the ability of AMPA to affect G α_{i1} in cortical neurons was also found in MIN6 cells. These data indicate that AMPA activated the G $_i$ proteins and that association between G α_{i1} and AMPA receptors was not initiated by G $_i$ proteins coupled to mGluRs on cortical cells, but was due to stimulation of the AMPA receptor.

Collectively, our data indicate that the AMPA receptor causes G $_i$ -protein activation in cortical neurons by a mechanism independent of its role as a Ca $^{2+}$ and Na $^{+}$ channel, and so it can operate with both ionotropic and metabotropic activity. In this regard, the AMPA receptor may be similar to the nicotinic acetylcholine receptor, which also appears capable of activating second messenger systems^{16,17}. Given that the predicted topology of AMPA-receptor subunits is different from that of typical G-protein-coupled receptors, it is unlikely that the $\alpha\beta\gamma$ heterotrimer docks directly with GluR1. It is possible that AMPA binding confers a conformational change to GluR1 that allows it to interact with an 'adaptor' protein, which in turn results in G $_i$ -protein binding and activation. Determining the exact nature of this interaction, its effect on downstream signalling, and ultimately its effect on neuronal function and behaviour, are issues currently being addressed. $\square$

Methods

Cell culture and fractionation. Cerebral cortical cultures prepared from 18-day-old fetal rats were plated on poly(L-lysine)-coated dishes^{18,19}. Cultures were used for experiments 10–13 days after plating. Unless stated, cultures were rinsed twice and placed in DMEM without serum immediately before experiments were initiated. MIN6 cells were cultured in DMEM containing 10% FBS as described^{14,15}. For Ca $^{2+}$ - and Na $^{+}$ -free conditions, the cells were rinsed twice and incubated in a solution containing (in mM) 120 choline chloride, 2 EGTA, 5.4 KCl, 20 glucose and 25 Tris-HCl, pH 7.4. Cells were lysed hypotonically and fractionated by sucrose gradient centrifugation^{20,21}.

ADP-ribosylation assay. Isolated membrane (66 μ g of protein) or cytosol (94 μ g of protein) was treated for 2 h at 30 °C with 40 μ l of ADP-ribosylation assay cocktail²² containing (in mM) 100 Tris-HCl, pH 7.5, 1 EGTA, 10 thymidine, 5 MgCl $_2$, 0.4 ATP, 0.4 GTP, 2.4 μ M nicotinamide adenine dinucleotide (NAD), 10 μ g ml $^{-1}$ PTX (pre-activated with 1 DTT at 37 °C for 30 min) and 0.0625 μ M [32 P]NAD (800 Ci mmol $^{-1}$, NEN). In the experiments of Figs 2 and 4, agents were added 3 min before initiating 2 h ADP-ribosylation assay. The reactions were terminated by the addition of Laemmli buffer²³. The samples were then separated on SDS-(7 or -10%) PAGE, transferred to immobilon-P and subjected to autoradiography. Unless indicated, 100 μ M AMPA or NMDA were used for ADP-ribosylation experiments. For G $\beta\gamma$ blockade experiments, isolated neuronal or MIN6 cell membranes, incubated with or without 25 ng of bovine brain G $\beta\gamma$ subunit (Chemicon) for 2 min, were then stimulated with 100 μ M AMPA for 3 min and ADP-ribosylation of G α_{i1} determined in the presence of both G $\beta\gamma$ and AMPA.

Determination of adenylate cyclase activity. Neuronal membranes in ADP-ribosylation cocktail (without PTX and NAD) were incubated with 5 μ M forskolin and 100 μ M AMPA for 10 min at 30 °C in the presence of 1 mM IBMX and then treated with 6% TCA. Levels of cAMP were determined by Biotrak EIA kit (Amersham).

Immunoprecipitation. For Fig. 3b, c, neurons were treated with 100 μ M AMPA, *trans*-(1S,3R)-ACPD or 1 unit thrombin for 3 min. Cultures were then rinsed twice with cold PBS containing 1 μ M leupeptin, 1 μ M pepstatin A and 1 mM PMSF, and 1 ml of lysis buffer (containing (in mM) 20 Tris-HCl, pH 7.4, 150 NaCl, 1 MgCl $_2$, 0.5% Triton X-100, 1 NaF, 20 Na $_3$ VO $_4$, 1 μ g ml $^{-1}$ leupeptin, 1 PMSF, 1 μ M pepstatin A and 1.0% bovine serum albumin) was added. The cells were scraped and homogenized. A 1:10 charcoal slurry was added and the extracts incubated for 1 h at 4 °C (ref. 19). Clarified supernatants were incubated with protein A-sepharose pre-absorbed with the anti-GluR1 antibody (UBI) overnight at 4 °C. The beads were collected, washed five times with buffer (containing (in mM) 20 Tris-HCl, pH 7.4, 150 NaCl, 20 EGTA and 0.5% Triton X-100) and the immunoprecipitate subjected to SDS-(10%) PAGE followed by electrotransfer to immobilon-P (Millipore). The Immobilon-P was

then analysed by western blot using antibodies against G α_{i1} or GluR1.

Western blot. Cell lysates from neurons or MIN6 cells containing equal amounts of proteins were separated and electrotransferred to Immobilon-P, which was blocked with 0.5% BSA and 5% powdered milk in PBS, and probed or reprobed with the indicated antibodies: Figs 1a, 2a with anti-G α_{i1} or anti-G α_s (Calbiochem); Fig. 3b, c with anti-G α_{i1} or anti-GluR1; Fig. 4a with anti-mGluR1-7 (Chemicon or gifts). The Immobilon-P was then washed five times with PBS containing 0.1% SDS, incubated with a horseradish peroxidase-conjugated goat anti-rabbit IgG, and washed a further five times. The bands were visualized by an Amersham ECL system.

RT-PCR. Total RNA was isolated from MIN6 cells or dissected rat retina using Tri Reagent RNA/DNA/protein isolation reagent (Molecular Research Center) and first-strand cDNA was made using Moloney murine leukaemia virus reverse transcriptase (Perkin Elmer Cetus) and primed with random hexamers. PCR reactions were performed in a final volume of 100 μ l using a Perkin Elmer PCR kit, AmpliTaq DNA polymerase and 50 pmol oligonucleotide primers (for mGluR6, TAGACGCTGTGTACGCCATTG and CGGTTGGTCTTGTTGAG-CAG). The PCR reaction was performed for 35 cycles (94 °C, 30 s; 60 °C, 30 s; 72 °C, 30 s) followed by 7 min final elongation at 72 °C. PCR products were electrophoresed and visualized by staining with ethidium bromide. GAPDH was also probed in both retina and MIN6 cells (primers, GGAGATTGTTGC-CATCAACGAC and ATGAGCCCTTCCACAATGCCAAAG) and served as a positive internal standard to assess levels of starting material and the efficiency of amplification.

Received 2 May; accepted 17 July 1997.

- Hollman, M. & Heinemann, S. Cloned glutamate receptors. *Annu. Rev. Neurosci.* **17**, 31–108 (1994).
- Seeburg, P. H. The TIPS/TINS lecture: the molecular biology of mammalian glutamate receptor channels. *Trends Pharmacol. Sci.* **14**, 297–303 (1993).
- Ghosh, A. & Greenberg, M. E. Calcium signaling in neurons: molecular mechanisms and cellular consequences. *Science* **268**, 238–247 (1995).
- Finkbeiner, S. & Greenberg, M. E. Ca $^{2+}$ -dependent route to ras: mechanisms for neuronal survival, differentiation, and plasticity. *Neuron* **16**, 233–236 (1996).
- Choi, D. W. Glutamate neurotoxicity and diseases of the nervous system. *Neuron* **1**, 623–634 (1988).
- Geiger, J. R. *et al.* Relative abundance of subunit mRNAs determines gating and Ca $^{2+}$ permeability of AMPA receptor in principal neurons and interneurons in rat CNS. *Neuron* **15**, 193–204 (1995).
- Carney, D. H. in *Thrombin Structure and Function* (ed. Berdinger, L. J.) 351–396 (Plenum, New York, 1992).
- Huff, R. M. & Neer, E. J. Subunit interactions of native and ADP-ribosylated $\alpha_{3\beta}$ and α_{41} , two guanine nucleotide-binding protein from bovine cerebral cortex. *J. Biol. Chem.* **261**, 1105–1110 (1986).
- Milligan, H. Techniques used in the identification and analysis of function and pertussis toxin-sensitive guanine nucleotide binding proteins. *Biochem. J.* **255**, 1–13 (1988).
- Katada, T., Oinuma, M. & Ui, M. Two guanine nucleotide-binding proteins in rat brain serving as the specific substrate of islet-activating protein, pertussis toxin. Interaction of the α -subunits with $\beta\gamma$ -subunits in development of their biological activities. *J. Biol. Chem.* **261**, 8182–8191 (1986).
- Brauner-Osborne, H. *et al.* A new highly selective metabotropic excitatory amino acid agonist: 2-amino-4-(3-hydroxy-5-methylisoxazol-4-yl)butyric acid. *J. Med. Chem.* **39**, 3188–3194 (1996).
- Nakajima, Y. *et al.* Molecular characterization of a novel retinal metabotropic glutamate receptor mGluR6 with a high agonist selectivity for L-2-amino-4-phosphonobutyrate. *J. Biol. Chem.* **268**, 11868–11873 (1993).
- Nomura, A. *et al.* Developmentally regulated postsynaptic localization of a metabotropic glutamate receptor in rat rod bipolar cells. *Cell* **77**, 361–369 (1994).
- Gono, T. *et al.* Functional neuronal receptors are expressed in the non-neuronal cell line MIN6. *J. Biol. Chem.* **269**, 16989–16992 (1994).
- Weaver, C. D. *et al.* Differential expression of glutamate receptor subtypes in rat pancreatic islets. *J. Biol. Chem.* **271**, 12977–12984 (1996).
- Giovannelli, A. *et al.* Acetylcholine induces voltage-independent increase of cytosolic calcium in mouse myotubes. *Proc. Natl Acad. Sci. USA* **88**, 10069–10073 (1991).
- Sorenson, E. M. & Gallagher, J. P. The membrane hyperpolarization of rat dorsolateral septal nucleus neurons is mediated by a novel nicotinic receptor. *J. Pharmacol. Exp. Ther.* **277**, 1733–1743 (1996).
- Durkin, J. P. *et al.* An early loss in membrane protein kinase C activity precedes the excitatory amino acid-induced death of primary cortical neurons. *J. Neurochem.* **66**, 951–962 (1996).
- Wang, Y. Z. & Durkin, J. P. α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid, but not N-methyl-D-aspartate, activates mitogen-activated protein kinase through G-protein $\beta\gamma$ subunits in rat cortical neurons. *J. Biol. Chem.* **270**, 22783–22787 (1995).
- Wang, Y. Z., Schramek, H. & Dunn, M. J. Cytosolic and nuclear mitogen-activated protein kinases are regulated by distinct mechanisms. *Exp. Cell Res.* **225**, 382–388 (1996).
- Chen, R.-H., Sarnecki, C. & Blenis, J. Nuclear localization and regulation of erk- and rsk-encoded protein kinases. *Mol. Cell. Biol.* **12**, 915–927 (1992).
- Kasahara, J. & Sugiyama, H. Inositol phospholipid metabolism in *Xenopus* oocytes mediated by endogenous G $_q$ and G $_i$ proteins. *FEBS Lett.* **355**, 41–44 (1994).
- Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).

Acknowledgements. We thank R. Tremblay, D. Song and S. MacLean for technical support and supplying neuronal and MIN6 cell cultures; R. Ball and A. Shapiro for the EIA; J. Conn for antibodies against mGluR4 and mGluR7; S. Nakanishi and R. Shigemoto for the antibody against mGluR6; S. Seino and J. Miyazaki for MIN6 cells; B. Chakravarthy for comments, and P. R. Walker for suggestions during the preparation of the manuscript. This work was partly supported by a Heart and Stroke Foundation of Ontario grant awarded under the OSCAR program (J.P.D. and Y.Z.W.).

Correspondence and requests for materials should be addressed to Y.Z.W. (e-mail: yizheng.wang@nrc.ca).

Prolonged photoresponses in transgenic mouse rods lacking arrestin

Jun Xu^{†§}, Robert L. Dodd^{‡§}, Clint L. Makino^{†‡}, Melvin I. Simon^{*}, Denis A. Baylor[‡] & Jeannie Chen^{*†}

^{*} Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

[‡] Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305, USA

[§] These authors contributed equally to this work.

[†] Present addresses: Cadus Pharmaceutical Corporation, 777 Old Saw Mill River Road, Tarrytown, New York 10591, USA (J.X.); Department of Ophthalmology, Harvard Medical School, Boston, Massachusetts 02114, USA (C.L.M.); and Mary D. Allen Laboratory for Vision Research, Doheny Eye Institute, Departments of Ophthalmology and Cell and Neurobiology, University of Southern California School of Medicine, Los Angeles, California 90033, USA (J.C.).

Arrestins are soluble cytoplasmic proteins that bind to G-protein-coupled receptors, thus switching off activation of the G protein and terminating the signalling pathway that triggers the cellular response^{1,2}. Although visual arrestin has been shown to quench the catalytic activity of photoexcited, phosphorylated rhodopsin in a reconstituted system³, its role in the intact rod cell remains unclear because phosphorylation alone reduces the catalytic activity of rhodopsin⁴⁻⁶. Here we have recorded photocurrents of rods from transgenic mice in which one or both copies of the arrestin gene were disrupted. Photoresponses were unaffected when arrestin expression was halved, indicating that arrestin binding is not rate limiting for recovery of the rod photoresponse,

as it is in *Drosophila*^{7,8}. With arrestin absent, the flash response displayed a rapid partial recovery followed by a prolonged final phase. This behaviour indicates that an arrestin-independent mechanism initiates the quench of rhodopsin's catalytic activity and that arrestin completes the quench. The intensity dependence of the photoresponse in rods lacking arrestin further suggests that, although arrestin is required for normal signal termination, it does not participate directly in light adaptation.

One or both copies of the arrestin gene were disrupted by homologous recombination in transgenic mice (Fig. 1a). The disruption is expected to prevent expression of both the full-length and the truncated splice variant of rod arrestin⁹. Retinas hemizygous for arrestin (+/-) contained approximately 30% of the normal amount of arrestin, whereas retinas from knockout mice (-/-) contained no arrestin (Fig. 1b). Immunoblots revealed that the levels of rod and cone transducin- α , transducin- β_1 , transducin- β_2 , phosphodiesterase- γ , phosducin and recoverin were similar in control (+/+), +/- and -/- retinas (data not shown), indicating that the absence of arrestin did not lead to substantial changes in the amounts of these visual transduction proteins.

Gross retinal morphology was little affected by the absence of rod arrestin. The number of photoreceptor nuclei in four-week-old -/- mice raised in cyclic light (12 h light : 12 h dark) was normal, but the rod outer segments were about 25% shorter than those of +/+ and +/- mice, and were somewhat disorganized (Fig. 2). These abnormalities were less evident in retinas from -/- mice born and raised in constant darkness (not shown). Furthermore, the rhodopsin content was lowered by as much as 20-fold in -/- retinas of four-week-old mice raised in cyclic light, but it was normal in -/- mice reared in darkness. There was thus a light-dependent reduction in the length and rhodopsin content of -/- outer segments as well as a disorganization of outer-segment structure. These effects are like

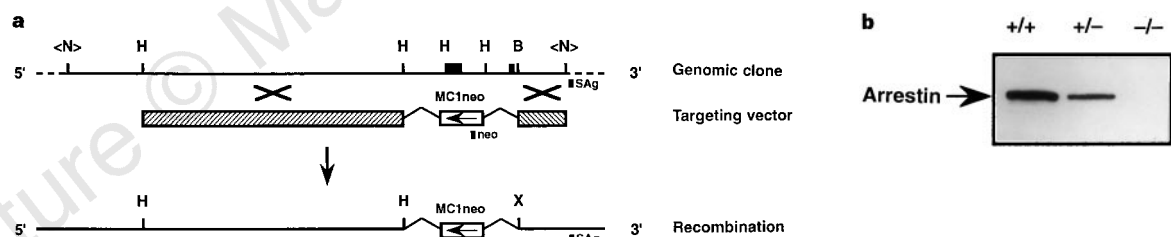


Figure 1 Targeted disruption of rod arrestin. **a**, Genomic structure, targeting vector and predicted structure of the disrupted arrestin locus. The targeting vector contained 7.2 kb of 5' and 1.3 kb of the 3' homologous sequence (hatched bars). The MC1neoP cassette replaced the 1.2-kb H-B fragment, which contained some 5' promoter elements as well as the first two exons (filled

boxes). H, *HindIII*; B, *BamHI*. **b**, Immunoblot analysis of rod arrestin from mice at 4 weeks of age probed with a monoclonal antibody raised against the 280-300 amino-acid region of bovine arrestin. Densitometric scans on blots from 5 sets of retinas gave the ratio of the arrestin content of +/- rods to that of +/+ as 0.33 ± 0.05 (mean $\pm$ s.d.); no arrestin was detectable in -/- rods.

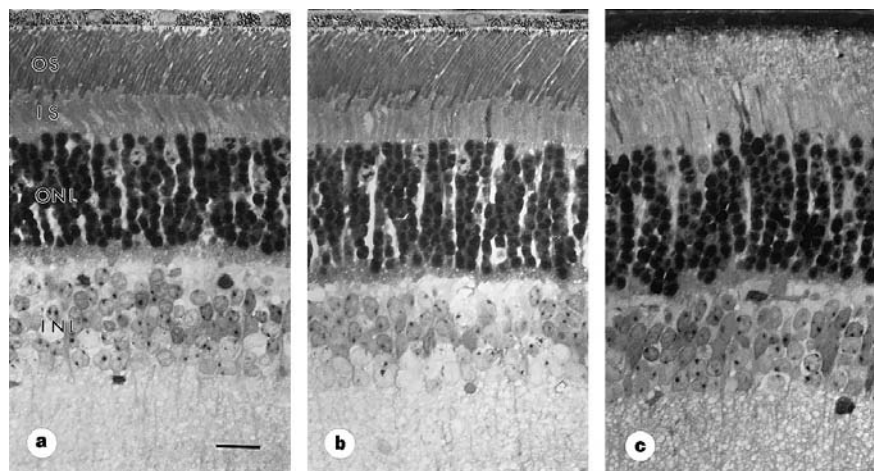


Figure 2 Morphology of the central retina of cyclic-light-reared control and arrestin knockout mice at 4 weeks of age. **a**, Control (+/+). **b**, Hemizygous knockout (+/-). **c**, Homozygous knockout (-/-). The outer nuclear layer, consisting of the photoreceptor nuclei, had similar thickness in all three retinas. The photoreceptor outer segments in -/- lacked the orderly arrangement characteristic of +/+ and +/- and were 25% shorter. OS, outer segments; IS, inner segments; ONL, outer nuclear layer; INL, inner nuclear layer. Scale bar, 16 μ m.

Table 1 Parameters of flash responses of control and transgenic rods

Cell type	t_p (ms)	t_i (ms)	a (pA)	i_0 (photons μm^{-2})	r_{max} (pA)
+/+	137 ± 2 (89)	240 ± 10 (89)	0.39 ± 0.04 (39)	49 ± 3 (77)	8.9 ± 0.3 (96)
+/-	120 ± 4 (7)	230 ± 37 (7)	0.44 ± 0.08 (5)	48 ± 2 (17)	8.5 ± 0.5 (17)
-/-	159 ± 9 (30)	$17,000 \pm 3,400$ (9)	0.32 ± 0.02 (21)	71 ± 7 (11)	6.4 ± 0.4 (18)
-/- dark-reared	139 ± 13 (7)	$10,700 \pm 932$ (5)		90 ± 77 (2)	8.9 ± 0.8 (9)

Values are mean $\pm$ s.e.m. with the number of cells studied in brackets; t_p is the time to peak of the dim flash response; t_i is the integration time of the dim flash response; a is the amplitude of the single photon response; i_0 is the flash strength at 500 nm that gave rise to a half-maximal response; and r_{max} is the saturating photoresponse amplitude.

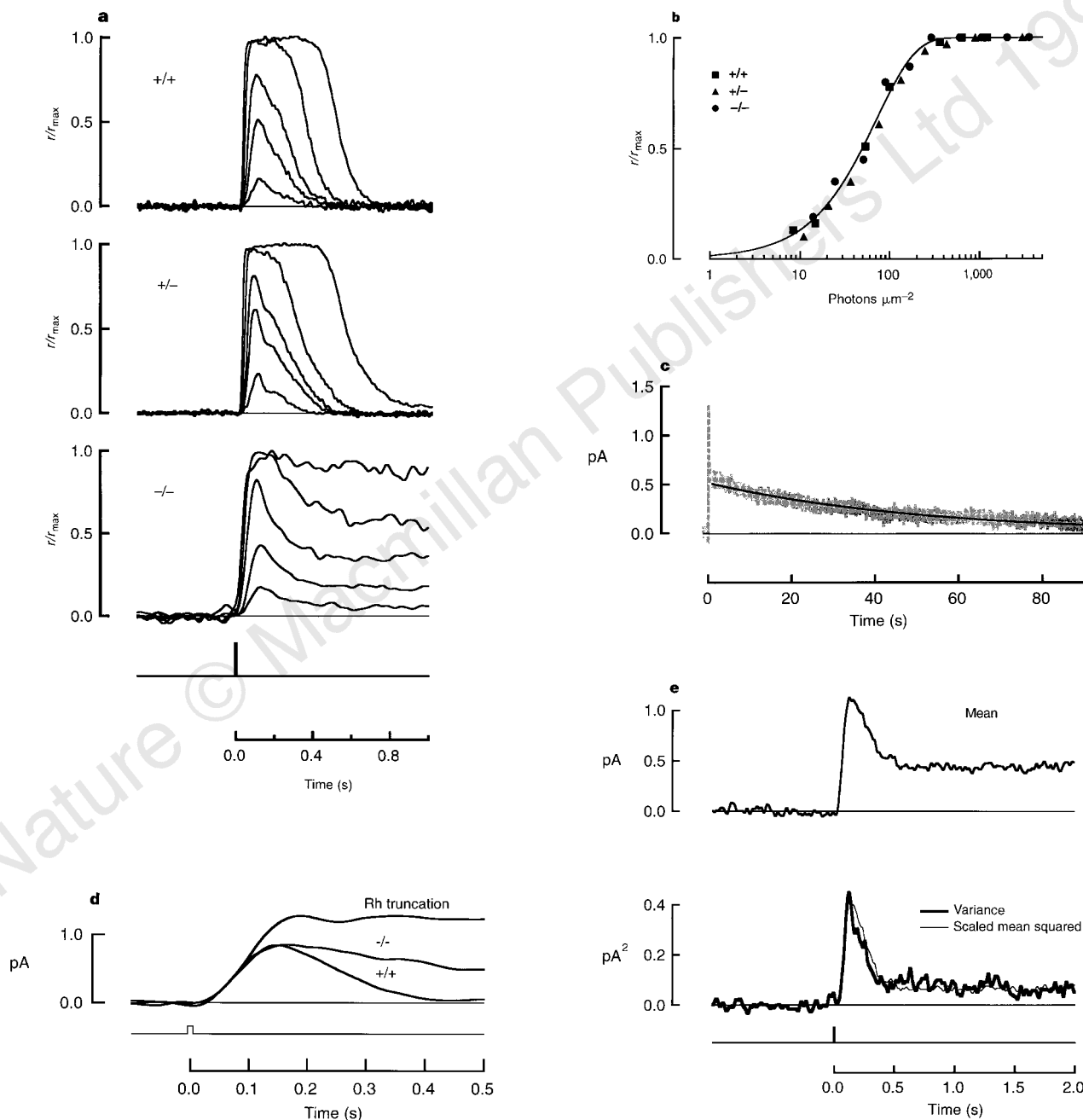


Figure 3 Rod flash responses. **a**, Averaged responses (r) of +/+ control (top), +/- (middle) and -/- (bottom) rods to flashes of increasing strength. Responses were normalized by the saturating response amplitudes (r_{max}): +/+, 18.9 pA; +/-, 14.1 pA; -/-, 7.6 pA. The largest -/- response was from a single trial. Flash strengths at 500 nm were: 8.3, 53.7, 99.4, 361 and 1,210 photons μm^{-2} for the +/+ rod; 20.9, 76, 133, 431 and 3,010 photons μm^{-2} for the +/- rod; and 14.1, 51.1, 166, 603 and 2,030 photons μm^{-2} for the -/- rod. Stimulus monitor trace is shown below the records. **b**, Stimulus-response relations. Peak amplitudes of responses in **a**, as well as points omitted in **a** for clarity. The continuous line is a saturating exponential, $r/r_{\text{max}} = 1 - \exp(-ki)$, where i is flash strength and k is a constant. **c**, Average of 63 dim flash responses from the -/- rod in **a** on an expanded time scale. The

black trace is an exponential with $\tau = 51$ s. **d**, Average single-photon response of a -/- rod compared with those observed¹⁶ when shut-off of rhodopsin was normal (+/+) or perturbed by prevention of C-terminal phosphorylation (Rh truncation). The -/- response has been scaled to have the same amplitude as the +/+ response. **e**, Ensemble variance increase during the dim flash response of a -/- rod. Top, mean of 54 responses to a flash estimated to cause an average of 3 photoisomerizations per trial. The ensemble variance increase (bottom, thick trace) coincided with the scaled square of the mean response (bottom, thin trace). From the ratio of the variance to the mean response, the single-photon response was estimated³⁰ as 0.4 pA. The flash strength was 13.5 photons μm^{-2} at 500 nm, and the estimated collecting area was 0.22 μm^2 .

those observed after chronic strong photic stimulation of the retina^{10,11}, and perhaps occur in $-/-$ mice because impaired rhodopsin shut-off greatly increases the duration and thus the effective gain of the light response (see below). The physiological experiments were performed on retinas from $-/-$ mice reared in the light-dark cycle or in complete darkness.

Flash responses from a control rod, a rod deficient in arrestin and a rod completely lacking it are compared in Fig. 3a. The responses of the $+/-$ rod were very similar to those of the control. Indeed, quantitative measures of flash sensitivity, dark current, and kinetics of the flash response were similar for the two populations of cells (Table 1).

Lowering the arrestin concentration should slow the rate of arrestin binding to photoisomerized rhodopsin, yet in $+/-$ rods, which contained only one-third the normal concentration of arrestin, the kinetics of the photoresponse were normal. This indicates that arrestin binding does not constitute a rate-limiting step in the recovery of the rod flash response. In contrast, the rate of recovery of the flash response in *Drosophila* photoreceptors varied linearly with the level of arrestin expression, indicating that arrestin binding is the rate-limiting step in recovery^{7,8}.

Complete deletion of arrestin left the rising phase of the flash response unaffected but strongly perturbed the recovery phase (Fig. 3a). Quantitative measures of flash sensitivity were similar in $-/-$

and $+/-$ rods, and the time to peak of the flash response was unaffected by arrestin deletion (Fig. 3a, b; Table 1). The rate of rise of the flash response in $-/-$ rods was very similar to that in control rods, consistent with normal transduction gain^{12,13}. The falling phase of the flash response in $-/-$ rods was abnormal, however (Fig. 3a, c), indicating that arrestin is essential for normal recovery. In $-/-$ rods the recovery of the flash response was biphasic: there was an initial rapid decline followed by a prolonged 'tail'. The biphasic recovery was present in the flash response of every $-/-$ rod studied (70 cells from 15 mice). The tail declined with a time constant of 42 ± 15 s (mean $\pm$ s.e.m., $n = 15$), similar to the time constant of thermal decay of rhodopsin's catalytically active form, metarhodopsin II, at 37°C (refs 14, 15), suggesting that thermal decay of metarhodopsin II may extinguish the catalytic activity of rhodopsin when arrestin is absent.

Before the tail, the amplitude of the dim flash response in $-/-$ mice fell to about half of its peak value; this decline occurred along the same time course as the fall in the normal response. These observations suggest that the catalytic activity of rhodopsin began to decline normally in $-/-$ rods, and that only the final process—arrestin binding—failed. The initial fall of catalytic activity in $-/-$ rods is presumably triggered by rhodopsin phosphorylation, because when the phosphorylation sites on rhodopsin were deleted the dim flash response rose to a larger amplitude that was sustained

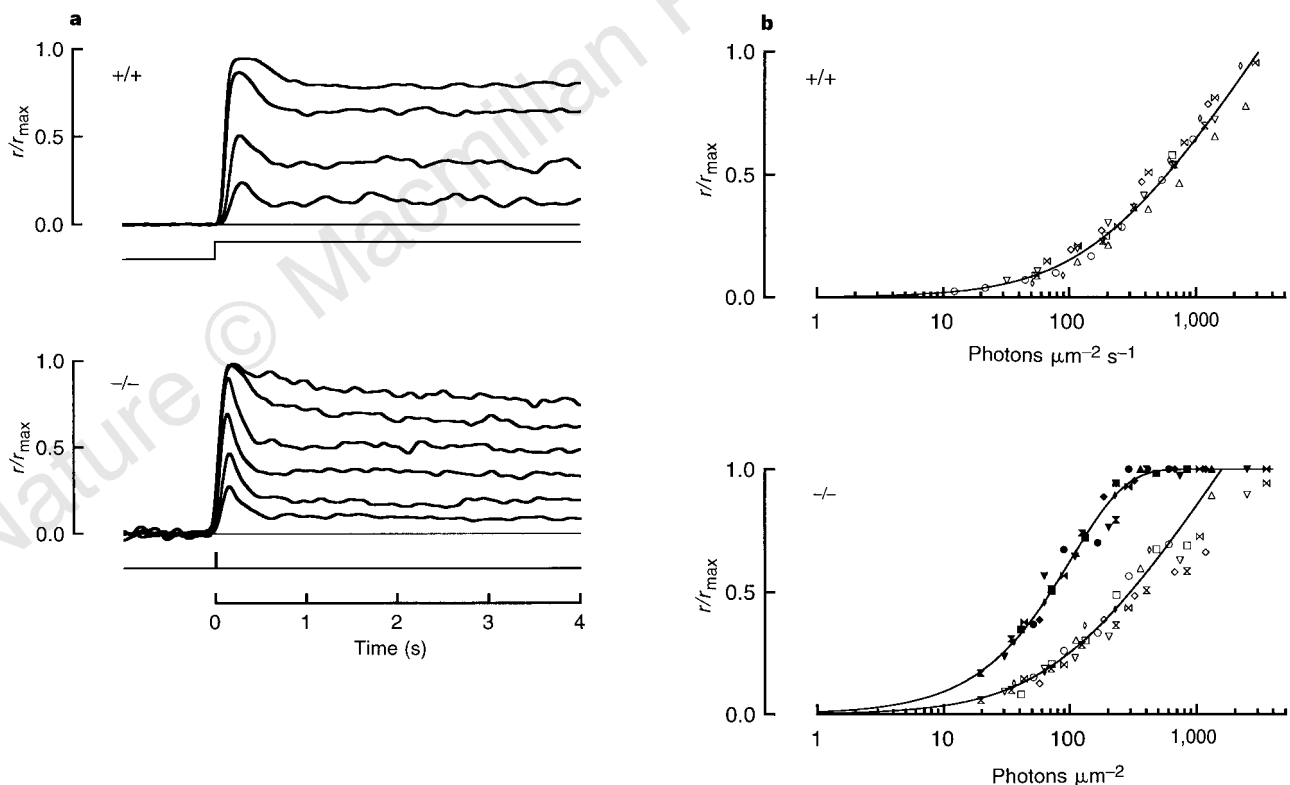


Figure 4 Lack of arrestin leaves the mechanism of light adaptation intact. **a**, Step response family from a control rod (top) and a flash response family from the $-/-$ rod of Fig. 3a (bottom). The bottom trace in each panel is a stimulus monitor. Saturating response amplitude for the control rod was 10 pA and the step intensities at 500 nm were 81.7, 297, 989 and 1,730 photons $\mu\text{m}^{-2} \text{s}^{-1}$. **b**, Collected stimulus-response relations for control rods to light steps (top, $n = 8$) and for $-/-$ rods to flashes (bottom, $n = 8$). Open symbols in each panel are plateau amplitudes measured 2 s after stimulus onset. Filled symbols in bottom panel are peak flash response amplitudes. Each set of points from a given cell was shifted on the abscissa by a factor of up to 5 to make the relation measured at the peak response coincide with the population mean at $r/r_{\max} = 0.5$. The continuous curve near the filled symbols is a saturating exponential with $k = 0.00976 \mu\text{m}^2 \text{photon}^{-1}$. The continuous curve in the top panel

was drawn according to the integral form of the Weber-Fechner relation²⁰, $r'/r_{\max} = I_0 S_0^D \ln(1 + I/I_0)$, where r' is the plateau amplitude of the response to a steady light of intensity I , S_0^D is the step sensitivity in darkness, and I_0 is the intensity that reduces the cell's step sensitivity to $0.5 S_0^D$. For the 8 rods, I_0 and S_0^D were 181 photons $\mu\text{m}^{-2} \text{s}^{-1}$ and $0.00185 \text{s} \mu\text{m}^2 \text{photon}^{-1}$, respectively. The curve in the top panel is redrawn near the open symbols in the bottom panel, where its position on the abscissa was found by assuming that, for flash and step responses of the same amplitude, $i = I t_i \alpha$, where i is the flash strength, I is the step intensity, t_i is the integration time of the $+/-$ flash response, taken as 257 ms, and α , taken as 2, is the factor by which the peak response amplitude is larger than the plateau in $-/-$. The curve should fit if the flash sensitivities and mechanisms of response compression and adaptation are identical in control and $-/-$ rods.

until shut-off¹⁶ (Fig. 3d). Apparently arrestin binding normally commences by 150 ms, when responses from +/+ and -/- rods begin to diverge (Fig. 3d).

Analysis of fluctuations in an ensemble of responses to dim flashes revealed that the single photon response of -/- rods consistently failed to shut off normally. The waveform of the time-dependent increase in ensemble variance resembled that of the mean response squared over the first two seconds after the flash (Fig. 3e). This would be expected if the number of quantal responses per flash fluctuated according to Poisson statistics and each quantal response had a stereotyped shape like that of the mean response (with its tail). The fact that the elementary response was consistently prolonged lends further support to the notion that arrestin binding is obligatory for the normal, rapid termination of rhodopsin's catalytic activity. In agreement with the effect of arrestin deletion we observed, phytic acid, which prevents the binding of arrestin to phosphorylated rhodopsin¹⁷, was found to slow the recovery of the flash response in isolated, dialysed rod outer segments from gecko¹⁸.

Comparison of responses from +/+ and -/- rods indicated that arrestin does not participate directly in light adaptation, which is the time-dependent, compressive nonlinearity that lowers the steady-state response amplitude and extends the range of light intensity in which transduction can operate successfully^{19,20}. In normal rods a bright step of light elicits a response that rises to a peak and then sags to a maintained response as adaptation occurs (Fig. 4a, top). Although the light response of -/- rods fails to shut off normally, a simple prediction should hold if the mechanisms that produce adaptation remain normal: the tails in the responses of -/- rods to light flashes (Fig. 4a, bottom) should be compressed in the same way as the steady-state responses of control rods to light steps. This follows because both types of maintained response are generated by persistent rhodopsin activity. The dependence of steady-state response amplitude on step intensity from experiments on eight control cells is shown in Fig. 4b (top). The continuous curve near the points is a fit of the relation expected to hold in a cell that adapts²⁰. The dependence of response amplitude on flash strength for -/- rods is shown in Fig. 4b at the peak response (filled symbols) and during the tail (open symbols); adaptation makes the tail relation flatter than the peak relation. The curve near the tail relation is that from Fig. 4b (top), placed on the abscissa at the position expected if the sole defect in -/- rods is the failure of the single photon response to shut off properly. The good fit suggests that arrestin does not itself participate in adaptation to background light. Therefore, although the mechanisms of adaptation were apparently normal in -/- rods, they failed to adapt normally because the long tails on the elementary responses superposed in steady light and drove the response into saturation at very low light intensity (data not shown).

In humans a naturally occurring frame-shift mutation is thought to produce non-functional arrestin²¹. Heterozygotes are asymptomatic, consistent with the lack of abnormality seen in the +/- mouse rods. Human homozygotes develop Oguchi disease, a form of congenital stationary night blindness. The reduced scotopic sensitivity and slowed dark adaptation in these patients may result from functional disturbances like those present in -/- transgenic mouse rods. □

Methods

Transgenic mice. Knockout mice were generated using standard procedures²². Transfected mouse embryonic stem (ES) cell clones were screened with primers *neo* 5'-CCATCTGTTCATGGCCGATCCC-3' and *Sag* 5'-CAGAGATCTGTGCTGCCACTTCCA-3' (small filled boxes in Fig. 1). Positive clones were incorporated into blastocysts to produce chimaeric mice, which were bred to homozygosity.

Western blot analysis. Retinas were isolated from four-week-old mice and homogenized in phosphate-buffered saline, pH 7.4. Samples were immunoblotted with S65-38, a monoclonal antibody raised against the 280–300 amino-

acid region of bovine arrestin²³ (gift from P. Hargrave, University of Florida). Rod and cone transducin- α were probed with antibodies obtained from M. Lochrie (California Institute of Technology). Other proteins were detected with antibodies that have been described previously: anti-rhodopsin (1D4) (ref. 24), anti-transducin- β_1 (ref. 25), anti-transducin- β_2 (ref. 25), anti-phosphodiesterase- γ (ref. 26), anti-phosphodiesterase- γ and anti-recoverin²⁸.

Histology. Eyecups were fixed overnight in 2.5% glutaraldehyde, 2% paraformaldehyde in 0.1 M phosphate buffer, pH 7.2, and processed in Epon. Sections 1 μ m thick from the central retina were stained with 0.5% methylene blue, 0.5% azure II in 0.5% borax.

Single-cell recording. Wild-type mice, aged 5–11 weeks, and transgenic mice aged 4–8 weeks, were reared in a 12 h light: 12 h dark cycle or in complete darkness. Mice were kept in darkness overnight, killed by CO₂ narcosis followed by decapitation, and the eyes enucleated under infrared illumination. Retinas were isolated into Leibovitz's L-15 medium (Gibco), cleared of vitreous humour and stored on ice. A small piece of retina was chopped finely in L-15 medium with DNase (type IV, Sigma) and the pieces were loaded into a recording chamber, which was perfused continuously with heated, gassed (95% O₂/5% CO₂) Locke's solution containing (in mM): 144 Na⁺, 3.6 K⁺, 1.2 Ca²⁺, 2.4 Mg²⁺, 123.3 Cl⁻, 10 HEPES, 20 HCO₃⁻, 0.02 EDTA, 10 glucose, BME vitamins, BME amino acids, 0.5 glutamate, 3 succinate, pH 7.4 at 36–38 °C. Using an infrared viewing system, a single rod outer segment was pulled into a suction electrode and the photocurrents recorded with a current-to-voltage converter (Axopatch 200, Axon Instruments). The electrode contained (in mM): 141.5 Na⁺, 3.6 K⁺, 1.2 Ca²⁺, 2.4 Mg²⁺, 140.8 Cl⁻, 10 HEPES, 0.02 EDTA, 10 glucose, pH 7.4. Recordings were low-pass filtered at 30 Hz (–3 dB) with an 8-pole Bessel filter and were digitized on-line at 400 or 100 Hz. Cells were stimulated with 10-ms flashes or steps of light at 500 nm. Control measurements include some results reported previously^{16,29}.

Received 3 April; accepted 2 July 1997.

- Wilson, C. J. & Applebury, M. L. Arresting G-protein coupled receptor activity. *Curr. Biol.* **3**, 683–686 (1993).
- Palczewski, K. Structure and functions of arrestins. *Protein Sci.* **3**, 1355–1361 (1994).
- Wilden, U., Hall, S. W. & Kuhn, H. Phosphodiesterase activation by photoexcited rhodopsin is quenched when rhodopsin is phosphorylated and binds the intrinsic 48-kDa protein of rod outer segments. *Proc. Natl Acad. Sci. USA* **83**, 1174–1178 (1986).
- Miller, J. L., Fox, D. A. & Litman, B. J. Amplification of phosphodiesterase activation is greatly reduced by rhodopsin phosphorylation. *Biochemistry* **25**, 4983–4988 (1986).
- Arshavsky, V. Y., Antoch, M. P., Dizhoor, A. M., Rakhilin, S. V. & Philippov, P. P. in *Retinal Proteins* (ed. Ovchinnikov, Y. A.) 497–503 (VNU, Utrecht, 1987).
- Wilden, U. Duration and amplitude of the light-induced cGMP hydrolysis in vertebrate photoreceptors are regulated by multiple phosphorylation of rhodopsin and by arrestin binding. *Biochemistry* **34**, 1446–1454 (1995).
- Dolph, P. J. et al. Arrestin function in inactivation of G protein-coupled receptor rhodopsin in vivo. *Science* **260**, 1910–1916 (1993).
- Ranganathan, R. & Stevens, C. F. Arrestin binding determines the rate of inactivation of the G protein-coupled receptor rhodopsin in vivo. *Cell* **81**, 841–848 (1995).
- Palczewski, K. & Smith, W. C. Splice variants of arrestins. *Exp. Eye Res.* **63**, 599–602 (1996).
- Penn, J. S. & Williams, T. P. Photostasis: regulation of daily photon-catch by rat retinas in response to various cyclic illuminances. *Exp. Eye Res.* **43**, 915–928 (1986).
- Schremmer, J. L. & Williams, T. P. Rod outer segment (ROS) renewal as a mechanism for adaptation to a new intensity environment. I. Rhodopsin levels and ROS length. *Exp. Eye Res.* **61**, 17–24 (1995).
- Lamb, T. D. & Pugh, E. N. Jr A quantitative account of the activation steps involved in phototransduction in amphibian photoreceptors. *J. Physiol. (Lond.)* **449**, 719–758 (1992).
- Pugh, E. N. Jr & Lamb, T. D. Amplification and kinetics of the activation steps in phototransduction. *Biochim. Biophys. Acta* **1141**, 111–149 (1993).
- Ebreys, T. G. The thermal decay of the intermediates of rhodopsin in situ. *Vis. Res.* **8**, 965–982 (1968).
- Cone, R. A. & Cobbs, W. H. Rhodopsin cycle in the living eye of the rat. *Nature* **221**, 820–822 (1969).
- Chen, J., Makino, C. L., Peachey, N. S., Baylor, D. A. & Simon, M. I. Mechanisms of rhodopsin inactivation in vivo as revealed by a COOH-terminal truncation mutant. *Science* **267**, 374–377 (1995).
- Palczewski, K., Pulvermiller, A., Buczylo, J., Gutmann, C. & Hofmann, K. P. Binding of inositol phosphates to arrestin. *FEBS Lett.* **295**, 195–199 (1991).
- Palczewski, K., Rispoli, G. & Detwiler, P. B. The influence of arrestin (48K protein) and rhodopsin kinase on visual transduction. *Neuron* **8**, 117–126 (1992).
- Baylor, D. A. & Hodgkin, A. L. Changes in time scale and sensitivity in turtle photoreceptors. *J. Physiol. (Lond.)* **242**, 729–758 (1974).
- Nakatani, K., Tamura, T. & Yau, K.-W. Light adaptation in retinal rods of the rabbit and two other nonprimate mammals. *J. Gen. Physiol.* **97**, 413–435 (1991).
- Fuchs, S. et al. A homozygous 1-base pair deletion in the arrestin gene is a frequent cause of Oguchi disease in Japanese. *Nature Genet.* **10**, 360–362 (1995).
- Ramirez-Solis, R., Davis, A. C. & Bradley, A. in *Methods in Enzymology: Guide to Techniques in Mouse Development* (eds Wassarman, P. M. & DePamphilis, M. L.) 855–878 (Academic, San Diego, 1993).
- Adamus, G., Machnicki, M. & Seigel, G. M. Apoptotic retinal cell death induced by anti-retroviral antibodies of cancer-associated retinopathy. *Invest. Ophthalmol. Vis. Sci.* **38**, 283–291 (1997).
- Molday, R. S. & MacKenzie, D. Monoclonal antibodies to rhodopsin: characterization, cross-reactivity, and application as structural probes. *Biochemistry* **22**, 653–660 (1983).
- Amatruda, T. T., Gautam, N., Fong, H. K., Northrup, J. K. & Simon, M. I. The 35- and 36-kDa beta subunits of GTP-binding regulatory proteins are products of separate genes. *J. Biol. Chem.* **263**, 5008–5011 (1988).
- Arshavsky, V. Y., Dumke, C. L. & Bownds, M. D. Noncatalytic cGMP-binding sites of amphibian rod cGMP phosphodiesterase control interaction with its inhibitory γ -subunits. *J. Biol. Chem.* **267**, 24501–24507 (1992).

27. Lee, R. H., Lieberman, B. S. & Lolley, R. N. Retinal accumulation of the phosphoducin/T β γ and transducin complexes in developing normal mice and in mice and dogs with inherited retinal degeneration. *Exp. Eye Res.* **51**, 325–333.
28. Dizhoor, A. M. *et al.* Recoverin: a calcium sensitive activator of retinal rod guanylate cyclase. *Science* **251**, 915–918 (1991).
29. Raport, C. J. *et al.* Downregulation of cyclic GMP phosphodiesterase induced by expression of GTPase-deficient cone transducin in mouse rod photoreceptors. *Invest. Ophthalmol. Vis. Sci.* **35**, 2932–2947 (1994).
30. Baylor, D. A., Lamb, T. D. & Yau, K.-W. Responses of retinal rods to single photons. *J. Physiol. (Lond.)* **288**, 613–634 (1979).

Acknowledgements. This work was supported by grants from the National Eye Institute (D.A.B.), the National Institute on Aging (M.J.S.), the National Institute of General Medical Sciences (R.L.D.), the Ruth and Milton Steinbach Fund (D.A.B., J.C.), The McKnight Endowment Fund for Neuroscience (D.A.B.), Research to Prevent Blindness (J.C., C.L.M.) and Lions' of Massachusetts (C.L.M.).

Correspondence and requests for materials should be addressed to D.A.B. (e-mail: dbaylor@leland.stanford.edu) or J.C. (e-mail: jeannie@hsc.usc.edu).

The exocytotic event in chromaffin cells revealed by patch amperometry

A. Albillos*, G. Dernick* $\ddagger$, H. Horstmann*, W. Almers*, G. Alvarez de Toledo† & M. Lindau* $\ddagger$

* Department of Molecular Cell Research, MPI f. Medical Research, D-69028 Heidelberg, Germany

† Department of Physiology & Biophysics, University of Seville, E-41009 Sevilla, Spain

In mast cells and granulocytes, exocytosis starts with the formation of a fusion pore^{1–3}. It has been suggested that neurotransmitters may be released through such a narrow pore without full fusion^{4,5}. However, owing to the small size of the secretory vesicles containing neurotransmitter, the properties of the fusion pore formed during Ca²⁺-dependent exocytosis and its role in transmitter release are still unknown. Here we investigate exocytosis of individual chromaffin granules by using cell-attached capacitance measurements^{3,6} combined with electrochemical detection of catecholamines^{7,8}, achieved by inserting a carbon-fibre electrode into the patch pipette. This allows the simultaneous determination of the opening of individual fusion pores and of the kinetics of catecholamine release from the same vesicle. We found that the fusion-pore diameter stays at <3 nm for a variable period, which can last for several seconds, before it expands. Transmitter is released much faster through this pore than in mast cells, generating a 'foot' signal⁸ which precedes the amperometric spike. Occasionally, the narrow pore forms only transiently and does not expand, allowing complete transmitter release without full fusion of the vesicle with the plasma membrane.

Patch amperometry combines capacitance measurements and amperometric detection of catecholamine in the cell-attached configuration by introducing a carbon-fibre electrode (CFE) into a patch pipette (Fig. 1a). Cell-attached capacitance measurements allow resolution of single fusion events on the size scale of synaptic vesicles^{3,6}. Vesicles fusing in the patch secrete their contents into the pipette. The CFE detecting released catecholamine must thus be placed inside the pipette. Figure 1b shows a cell at the beginning of the experiment with a patch pipette containing a CFE positioned ~6–7 μ m from the pipette tip. After pipette attachment, exocytosis of chromaffin granules was indicated by amperometric signals associated with capacitance steps: part of the recording is shown in Fig. 1d. At the end of the experiment, the morphology of the cell had changed (Fig. 1c), indicating that massive exocytosis had occurred not only in the patch but all over the cell surface. This cell was pretreated with phorbol myristate acetate (PMA) to increase the pool of readily releasable vesicles⁹.

After sealing the patch pipette to the plasma membrane, exocytotic events occurred spontaneously in active patches. After steps had ceased, patch depolarization did not induce further capacitance steps. Similarly, in patches showing no spontaneous activity, depolarization did not stimulate any capacitance step or amperometric transient. Thus, sealing of the pipette is sufficient to stimulate exocytosis of all releasable granules in the patch. No exocytotic events were observed when the bath and pipette solution contained no Ca²⁺ and 1 mM EGTA (14 patches), indicating that exocytosis is

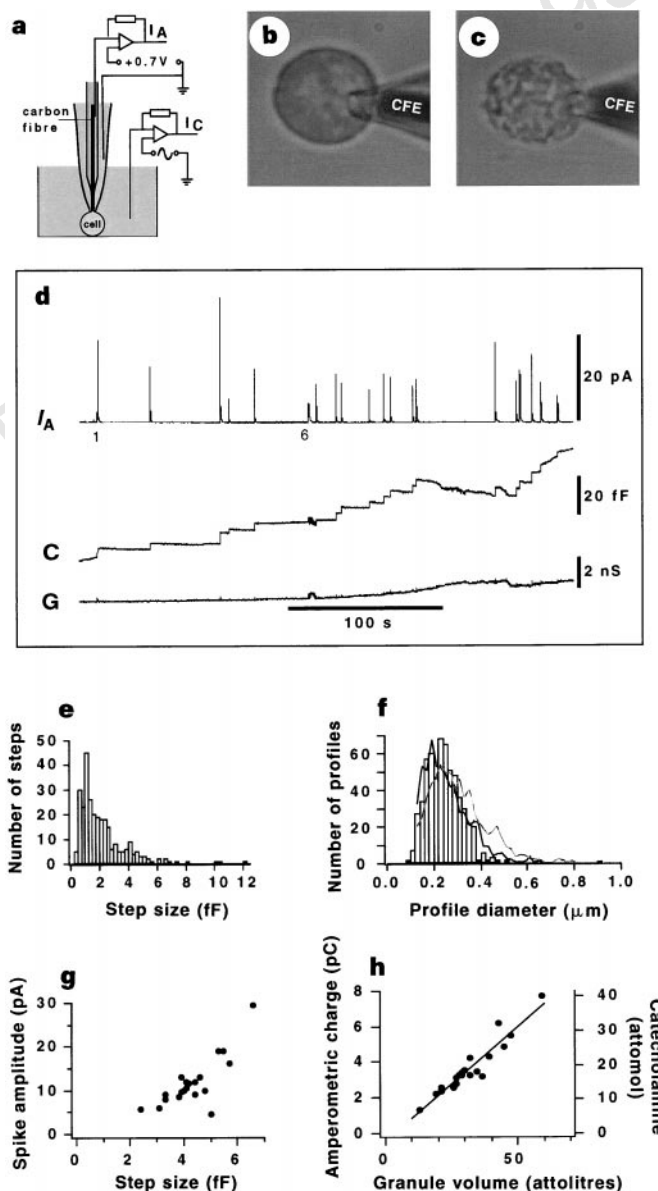


Figure 1 a, Arrangement of a CFE inside a patch pipette. *I_A*, Amperometric current; *I_C*, sine wave current used to measure capacitance changes. b, c, Chromaffin cell with attached patch pipette containing CFE at the beginning (b) and end (c) of the experiment. d, Recording from this cell shows amperometric transients (top), associated capacitance steps (middle), and conductance trace (bottom) using a 20-kHz, 50-mV sine wave. C is capacitance trace, G is conductance trace. Events 1 and 6 (labelled) are analysed in more detail in Figs 2, 3. e, Capacitance step size distribution obtained in experiments without PMA preincubation. f, Bars: frequency distribution of granule profile diameter from electron microscopy of resting chromaffin cells. Lines: profile size distributions expected from the conversion of capacitance step sizes, assuming specific capacitances of 9 fF μ m² (black) or 6 fF μ m² (grey). g, Relation between amperometric current amplitude and capacitance step size. h, Relation between catecholamine content and granule volume.

$\ddagger$ Present address: School of Applied and Engineering Physics, Cornell University, Ithaca, New York 14853-2501, USA.

due to Ca^{2+} influx rather than to increased membrane tension, which stimulates transmitter release in the frog neuromuscular junction independently of calcium^{10,11}. Mechanical stimulation during the sealing process is probably mediated by opening of chloride channels, leading to depolarization and Ca^{2+} influx¹².

Capacitance steps associated with amperometric spikes ranged from 0.3 to 12 fF, with a main peak at ~ 1 fF (Fig. 1e). Average step size was 2.1 ± 1.7 fF (s.d.; $n = 286$), similar to previous estimates^{6,13}. For comparison with electron micrographs, we converted the capacitance step sizes into the distribution of profile sizes that would be expected in thin sections, assuming integer values of $6\text{--}11 \text{ fF } \mu\text{m}^{-2}$ for the specific capacitance of the vesicle membrane^{14,15}. The best agreement with our morphometric data (Fig. 1f, histogram bars) was obtained assuming $9 \text{ fF } \mu\text{m}^{-2}$ (black line).

The average charge of amperometric transients was 2.2 ± 2.1 pC

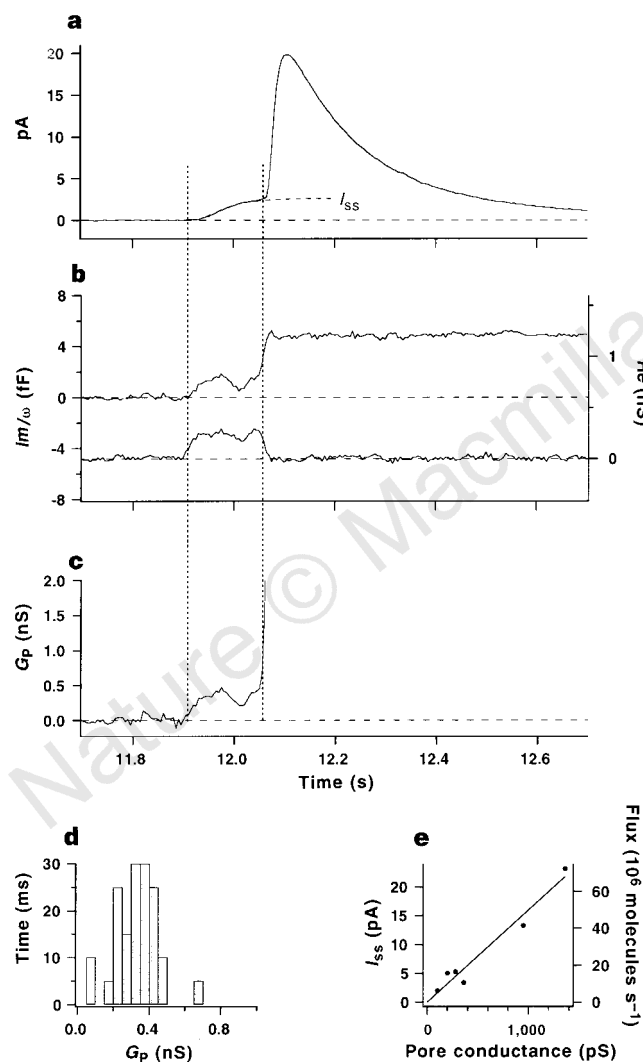


Figure 2 First exocytotic event shown in Fig. 1d on an expanded scale, showing amperometric current (**a**), imaginary part of the admittance (**b**, upper trace), and real part of admittance (**b**, lower trace) after baseline subtraction. **c**, Fusion pore conductance. **d**, Histogram of conductance levels during the foot (between vertical dashed lines in **a**–**c**). Long foot signals approached an apparent steady state I_{ss} , which was obtained by fitting the function $I_A = I_{ss} - I_0 \exp(-kt)$ to the foot signal (dashed line in **a**). **e**, Proportional relation between I_{ss} and average fusion pore conductance G_P during the foot (0.016 pA pS^{-1}).

(s.d.; $n = 172$) which is in the range reported^{16–18}. The amplitude of individual release transients is correlated with the size of the associated capacitance steps (Fig. 1g). To determine the catecholamine concentration inside the vesicles, the catecholamine content was calculated from the charge of the amperometric transient, assuming a transfer of two electrons per molecule¹⁹, and step sizes were converted into granule volumes. The catecholamine content is proportional to granule volume (Fig. 1h), indicating that in this cell vesicles of different sizes had a constant catecholamine concentration of 0.7 M. The concentration obtained from 172 exocytotic events in 9 patches was 0.71 ± 0.02 M. In individual patches, the slope of the regression line was 0.45–1.14 M. In PMA-treated cells, both catecholamine content and step sizes were generally larger, suggesting that PMA-treated cells preferentially secrete larger granules or perform compound exocytosis. However, catecholamine concentrations were similar (0.96 ± 0.02 M).

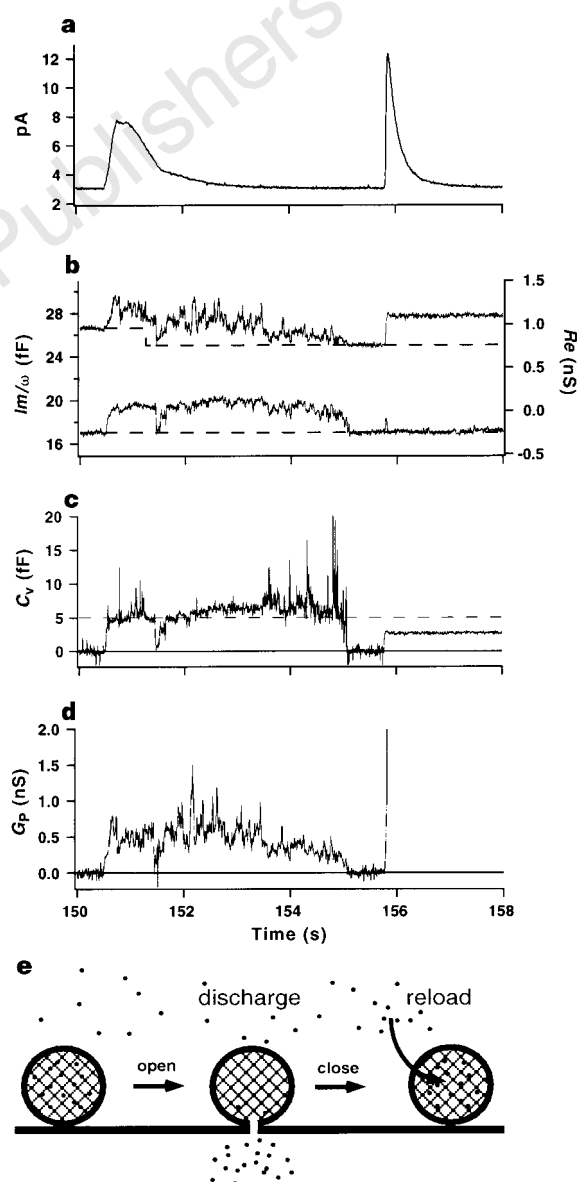


Figure 3 Event 6 shown in Fig. 1d on an expanded timescale, showing amperometric trace (**a**), imaginary part of admittance (**b**, upper trace), and real part of admittance (**b**, lower trace). After appropriate baseline subtraction (dashed lines) C_V (**c**) and G_P (**d**) were calculated. C_V values are more noisy when G_P is small. **e**, Transient fusion pore opening allows complete transmitter release without loss of matrix proteins, allowing full reloading.

The time course of all amperometric transients of Fig. 1d (except the sixth) was very similar, with a half-width of 190 ± 20 ms (s.d.; $n = 19$). The long duration of the transients in this experiment is presumably due largely to the diffusion across the 6–7- μ m distance between the pipette tip and the FE^{8,20–22}. In other experiments in which the CFE could be positioned within 3 μ m from the pipette tip, transients with half-width <20 ms were obtained (data not shown).

Many amperometric spikes were preceded by a 'foot' signal as previously described⁸. In patches with short diffusion distances, indicated by transients of half-width <20 ms, 52% of the spikes had a detectable foot, similar to the 50–70% reported for whole-cell amperometry^{8,18}. In experiments with larger diffusion distances, only very long foot signals were detectable. Figure 2 shows the first event of Fig. 1d on an expanded scale. During the foot (Fig. 2a), the imaginary part (Fig. 2b, upper trace), usually indicating capacitance, is much lower than the final level and a transient increase is seen in the real part (Fig. 2b, lower trace). This phenomenon indicates that the fusion pore connecting the interior of the granule to the extracellular medium had a low conductance. From the two traces, the fusion pore conductance, G_p , was calculated (Fig. 2c). During the foot, G_p varied between 200 and 500 pS; Fig. 2d shows the histogram of conductance levels during the foot. The mean value of G_p between the vertical dashed lines was 360 pS. In 39 events that gave a clear foot signal, the mean pore conductance during the foot was 515 ± 259 pS (s.d.; $n = 39$), corresponding to fusion pore diameters of 2.4 ± 0.7 nm.

In some cases (as in Fig. 2), the fusion pore was stable for long enough for a steady-state current to be estimated for the foot in the amperometric trace where release through the fusion pore equals consumption by the CFE. As in rat peritoneal mast cells²³, the release rate measured as the amperometric current is proportional to G_p (Fig. 2e). Linear regression gives a slope of 0.016 pA pS⁻¹, corresponding to a flux of 5×10^7 molecules per second per nS of pore conductance. This is at least 10 times faster than serotonin release from mast-cell granules through pores of a similar size²³, presumably because of a much higher concentration of free transmitter. From the flux per nS, we estimate a free catecholamine concentration of ~250 mM inside chromaffin granules. As the catecholamine molecules are positively charged, the flux of counterions into the vesicle may also be rate-limiting²⁴, so 250 mM should be considered a lower limit for the free concentration.

Figure 3 shows event 6 of Fig. 1d on an expanded timescale. This amperometric transient has a small amplitude and unusual shape (Fig. 3a). Imaginary and real parts of the admittance (Fig. 3b) show correlated fluctuations of similar size, which is typical for a small fusion pore, followed by a return to the baseline. From these traces we estimate a vesicle capacitance of 5 fF (Fig. 3c, dashed line). This vesicle was transiently connected to the extracellular space for 4.6 s. During this transient fusion, the pore conductance fluctuated around 500 pS (Fig. 3d), a value similar to that observed during foot signals (Fig. 2c). The transient formation of such a fusion pore explains the so-called "stand-alone foot signals" recorded from chromaffin cells¹⁸ with whole cell amperometry, as well as similar recordings from leech neurons²⁵.

The charge from integrating this transient was 4.34 pC, corresponding to 1.4×10^7 molecules, or 23 attomol. The step size of 5 fF corresponds to a vesicle diameter of 420 nm, giving a volume of 39 attolitres and a catecholamine concentration of 0.6 M for this vesicle. The corresponding point in Fig. 1h is close to the regression line. Thus, virtually all catecholamine of this vesicle was released during the transient fusion pore opening. This agrees well with the observation in whole-cell amperometry that the average release during stand-alone foot signals was 77% of that measured in spike signals¹⁸. Most of the catecholamine of this large vesicle was released during the first second (Fig. 3a). As the volume of a synaptic vesicle is 1,000 times smaller, it would discharge most of its contents in less

than 1 ms through a similar fusion pore.

The next capacitance step is smaller (3 fF; Fig. 3c) and is associated with full catecholamine release (0.73 M). As catecholamine uptake by chromaffin granules is very slow^{26,27}, this event reflects exocytosis of a different smaller vesicle, with full release. During the remaining part of the recording, there was no capacitance step with or without significantly reduced catecholamine release, indicating that the vesicle releasing its catecholamine during a transient pore opening never fused completely with the plasma membrane. In our experiments, 65 of 458 vesicles formed transient fusion pores associated with significant transmitter release, similar to the relative frequency of stand-alone foot signals¹⁸. In 27 events, the noise was small enough to estimate the charge and capacitance. This analysis showed that at least 20 of the 27 events (74%) were true kiss-and-run events in which the vesicle did not fuse completely within the next 3–10 min. Assuming the same ratio for the more noisy events, we estimate that there were at least 48 kiss-and-run events, corresponding to more than 10% of all exocytotic events. This contrasts with previous results from mast cells, in which the transient events were always followed by complete fusion of the vesicle with the plasma membrane and only a very small fraction of secretory products was released during transient fusion pore openings²³.

Simultaneous measurements of the time course of fusion and release from single vesicles have previously only been reported for giant granules with capacitance >100 fF (ref. 23). We developed a new technique called patch amperometry which extends these measurements to small vesicles with capacitance as small as 0.1 fF; using this method, we found that in single chromaffin cells all granules, regardless of size, contain approximately the same concentration of catecholamines. The average concentration is ~0.7–0.9 M, close to the 0.6 M obtained by biochemical analysis²⁸. We showed that the amperometric foot reflects release of $\sim 2.5 \times 10^7$ molecules per second through fusion pores with conductances of around 500 pS, a rate much higher than had previously been reported in mast cells. Such a pore may be formed transiently, allowing complete release of the granular contents without undergoing full fusion. As a fusion pore smaller than 3 nm is unlikely to allow permeation of proteins as large as chromogranin A (relative molecular mass 77K), this mechanism would allow a vesicle to discharge all its catecholamine without losing the proteins of the granular matrix (Fig. 3e). After closing of the fusion pore, the vesicle may be reused and reloaded with catecholamines. The matrix proteins are required to store catecholamines at high concentration. Therefore, full reloading would not be possible if they were exocytosed and granules would have to be synthesized *de novo* at the *trans*-Golgi network. □

Methods

Chromaffin cells were prepared and cultured as described²⁹. Recordings were made on days 1–7 in culture. The bath solution contained (in mM) 140 NaCl, 5 KCl, 5 CaCl₂, 1 MgCl₂, 10 HEPES/NaOH, 20 glucose (pH 7.3). To increase the pool of readily releasable vesicles, cells were preincubated in the presence of 100 nM phorbol myristate acetate (PMA) for 6 min before sealing in some experiments⁹. PMA was added from a 1 mM stock solution in DMSO. This pretreatment increased the fraction of patches undergoing exocytotic events from 24% (13 out of 55 patches) without PMA pretreatment, to 79% (19 out of 24 patches) with PMA pretreatment. PMA-treated cells had larger capacitance steps (3.1 ± 2.5 fF, mean $\pm$ s.d.; $n = 126$) with a larger amperometric charge (5.2 ± 8.2 pC, mean $\pm$ s.d.; $n = 82$), facilitating the analysis of single events. Unless stated otherwise, data are from cells without PMA treatment. The pipette solution contained 50 NaCl, 100 TEA-Cl, 5 KCl, 5 CaCl₂, 1 MgCl₂, 10 HEPES/NaOH (pH 7.3). All experiments were done at room temperature.

Changes in membrane capacitance and catecholamine release were recorded simultaneously in the cell-attached configuration with a patch pipette containing a CFE. A pipette holder containing two Ag/AgCl electrodes was designed. The pipette electrode immersed into the pipette solution was

connected to ground. The other electrode made contact with the CFE via a 3 M KCl solution; the holder consisted of two parts screwed onto each other. The pipette could thus be adjusted until the CFE was as close to the pipette tip as possible. The pipette resistance increased less than twofold during the adjustment.

The CFE was prepared by cannulating a 5- μ m-diameter carbon fibre into polyethylene tubing (diameter: outer, 0.8 mm; inner, 0.4 mm). The tip of the CFE was pulled as described²², except that the narrow part at the tip was longer (about 500 μ m) to allow positioning inside the patch pipette close to the pipette tip. The polyethylene tubing was filled with 3 M KCl and connected to the Ag/AgCl wire which was held at +650–700 mV. Amperometric currents were recorded using a home-made amplifier and were filtered with either a 10-ms time constant or an 8-pole Bessel filter set to 100 Hz. Pipettes were pulled with a programmable puller (P97, Sutter Instruments), coated with sticky wax (Kerr) and heat-polished. Tip sizes were 2–5 μ m and pipette resistances were 1–3 M Ω .

The bath electrode was connected to the input of the patch clamp amplifier head stage (EPC7, HEKA-Elektronik). Changes of patch admittance were measured as described^{3,6} with a lock-in amplifier (SR 830, Stanford Research Systems) using sine-wave amplitudes of 20–50 mV (r.m.s.) and frequencies of 8 or 20 kHz, with output filters set to 1 or 3 ms, 24 dB. When necessary, recordings were digitally filtered with a 50-Hz notch filter to remove line-frequency noise.

To convert the capacitance step size distribution into the expected profile size distributions, the steps were converted into granule diameters, assuming spherical geometry and integer values of 6–11 fF μ m² for the specific capacitance. The resulting granule size distribution was then converted into a profile size distribution taking into account the random sectioning, as described in equation 5.5 of ref. 30. The amplitudes of these distributions were adjusted by an unweighted least-squares fit to the profile size histogram using the range of profile diameters >160 nm. The best fit was obtained for 9 fF μ m².

Fusion pore openings were analysed as described³. Vesicle capacitance (C_v) and fusion pore conductance (G_p) were calculated from the real (Re) and imaginary (Im) parts of the admittance after baseline subtraction: $C_v = [(Re^2 + Im^2)/Im]/\omega$ and $G_p = (Re^2 + Im^2)/Re$. The pore diameter was estimated from G_p as described^{1,2}, assuming a cylindrical pore 15 nm long and $\sigma = 15 \text{ mS cm}^{-1}$ for the conductivity of the solution.

The relation between fusion pore conductance and rate of transmitter release was determined for those foot signals in which the fusion pore was stable for long enough for the amperometric current to approach a steady-state value. In these cases, the amperometric current during the foot was fitted by the function $I_A = I_{ss} - I_0 \exp(-t/\tau)$ after baseline subtraction (Fig. 2a). The average pore conductance was determined during the time between pore opening and rapid pore expansion, as indicated by vertical dashed lines in Fig. 2. Each data point in Fig. 2e is derived from a different fusion event. Assuming a pore with cross-sectional area a and length b , the flux j of molecules per second is $j = Da/b\Delta C$ (where $D = 6 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$ is the diffusion coefficient²² and ΔC is the concentration gradient). With $G_p = \sigma a/b$, the relation $\Delta C = j\sigma/(G_p D)$ is obtained. Owing to the dilution of catecholamine in the pipette, this is approximately equal to the free catecholamine concentration inside a chromaffin granule.

Received 23 June; accepted 24 July 1997.

1. Breckenridge, L. J. & Almers, W. Currents through the fusion pore that forms during exocytosis of a secretory vesicle. *Nature* **328**, 814–817 (1987).
2. Spruce, A. E., Breckenridge, L. J., Lee, A. K. & Almers, W. Properties of the fusion pore that forms during exocytosis of a mast cell secretory vesicles. *Neuron* **4**, 643–654 (1990).
3. Lollike, K., Borregaard, N. & Lindau, M. The exocytotic fusion pore of small granules has a conductance similar to an ion channel. *J. Cell Biol.* **129**, 99–104 (1995).
4. Almers, W. & Tse, F. W. Transmitter release from synapses: Does a preassembled fusion pore initiate exocytosis? *Neuron* **4**, 813–818 (1990).
5. Neher, E. Secretion without full fusion. *Nature* **363**, 497–498 (1993).
6. Neher, E. & Marty, A. Discrete changes of cell membrane capacitance observed under conditions of enhanced secretion in bovine adrenal chromaffin cells. *Proc. Natl Acad. Sci. USA* **79**, 6712–6716 (1982).
7. Wightman, R. M. et al. Temporally resolved catecholamine spikes correspond to single vesicle release for individual chromaffin cells. *Proc. Natl Acad. Sci. USA* **88**, 10754–10758 (1991).
8. Chow, R. H., Rüden, L. v. & Neher, E. Delay in vesicle fusion revealed by electrochemical monitoring of single secretory events in adrenal chromaffin cells. *Nature* **356**, 60–63 (1992).
9. Gillis, K. D., Mößner, R. & Neher, E. Protein kinase C enhances exocytosis from chromaffin cells by increasing the size of the readily releasable pool of secretory granules. *Neuron* **16**, 1209–1220 (1996).
10. Chen, B.-M. & Grinell, A. D. Integrins and modulation of transmitter release from motor nerve terminals by stretch. *Science* **269**, 1578–1580 (1995).
11. Chen, B.-M. & Grinell, A. D. Kinetics, Ca^{2+} -dependence, and biophysical properties of integrin-

- mediated mechanical modulation of transmitter release from frog motor nerve terminals. *J. Neurosci.* **17**, 904–916 (1997).
12. Moser, T., Chow, R. H. & Neher, E. Swelling-induced catecholamine secretion recorded from single chromaffin cells. *Pflügers Arch. Eur. J. Physiol.* **431**, 196–203 (1995).
13. Chow, R. H., Klingauf, J., Heinemann, C., Zucker, R. S. & Neher, E. Mechanisms determining the time course of secretion in neuroendocrine cells. *Neuron* **16**, 369–376 (1996).
14. Breckenridge, L. J. & Almers, W. Final steps in exocytosis observed in a cell with giant secretory granules. *Proc. Natl Acad. Sci. USA* **84**, 1945–1949 (1987).
15. Zimmerberg, J., Curran, M., Cohen, E. S. & Brodwick, M. Simultaneous electrical and optical measurements show that membrane fusion precedes secretory granule swelling during exocytosis of beige mouse mast cells. *Proc. Natl Acad. Sci. USA* **84**, 1585–1589 (1987).
16. Walker, A., Glavinovic, M. I. & Trifaró, J.-M. Temperature dependence of release of vesicular content in bovine chromaffin cells. *Pflügers Arch. Eur. J. Physiol.* **432**, 885–892 (1996).
17. Jankowski, J. A., Schroeder, T. J., Ciolkowski, E. L. & Wightman, R. M. Temporal characteristics of quantal secretion of catecholamines from adrenal medullary cells. *J. Biol. Chem.* **268**, 14694–14700 (1993).
18. Zhou, Z., Misler, S. & Chow, R. H. Rapid fluctuations in transmitter release from single vesicles in bovine adrenal chromaffin cells. *Biophys. J.* **70**, 1543–1552 (1996).
19. Baur, J. E., Kristensen, E. W., May, L. J., Wiedemann, D. J. & Wightman, R. J. Fast-scan voltammetry of biogenic amines. *Anal. Chem.* **60**, 1268–1272 (1988).
20. Walker, A., Glavinovic, M. I. & Trifaró, J.-M. Time course of release of content of single vesicles in bovine chromaffin cells. *Pflügers Arch. Eur. J. Physiol.* **431**, 729–735 (1996).
21. Wightman, R. M., Schroeder, T. J., Finnegan, J. M., Ciolkowski, E. L. & Pihel, K. Time course of release of catecholamines from individual vesicles during exocytosis at adrenal medullary cells. *Biophys. J.* **68**, 383–390 (1995).
22. Chow, R. H. & Rüden, L. v. in *Single Channel Recording* (eds Sakmann, B. & Neher, E.) 245–275 (Plenum, New York, 1995).
23. Alvarez de Toledo, G., Fernández-Chacón, R. & Fernandez, J. M. Release of secretory products during transient vesicle fusion. *Nature* **363**, 554–558 (1993).
24. Rahamimoff, R. & Fernandez, J. M. Pre- and postfusion regulation of transmitter release. *Neuron* **18**, 17–27 (1997).
25. Bruns, D. & Jahn, R. Real-time measurement of transmitter release for single synaptic vesicles. *Nature* **377**, 62–65 (1997).
26. Gasnier, B., Scherman, D. & Henry, J.-P. Inactivation of the catecholamine transporter during the preparation of chromaffin granule 'ghosts'. *FEBS Lett.* **222**, 215–219 (1987).
27. Scherman, D. & Boschi, G. Time required for transmitter accumulation inside monoaminergic storage vesicles differs in peripheral and central systems. *Neuroscience* **27**, 1029–1035 (1988).
28. Phillips, J. H. Dynamic aspects of chromaffin granule structure. *Neuroscience* **7**, 1595–1609 (1982).
29. Parsons, T. D., Coorsen, J. R., Horstmann, H. & Almers, W. Docked granules, the exocytic burst, and the need for ATP hydrolysis in endocrine cells. *Neurons* **15**, 1085–1096 (1995).
30. Bach, G. in *Quantitative Methoden der Morphologie* (eds Weibel, E. R. & Elias, H.) 23–45 (Springer, Berlin, 1967).

Acknowledgements. We thank J. Pauli and the staff of the mechanical workshop for their contributions to the development and production of the adjustable pipette holder; J. Tritthard for the development of the amperometric amplifier; I. Wunderlich and R. Hinz-Herkommer for technical assistance; and E. Neher for his comments and encouragement. This work was supported by grants from the Deutsche Forschungsgemeinschaft to M.L. and from the Spanish Ministerio de Educación y Cultura (DGICYT) to G.A.T. A.A. was supported by an FPI fellowship from DGICYT.

Correspondence and requests for materials should be addressed to M.L. (e-mail: ml95@cornell.edu).

Circadian oscillation of a mammalian homologue of the *Drosophila period* gene

Hajime Tei*, Hitoshi Okamura†, Yasufumi Shigeyoshi†, Chiaki Fukuhara‡, Ritsuko Ozawa*, Matsumi Hirose* & Yoshiyuki Sakaki*

* Human Genome Center, Institute of Medical Science, University of Tokyo, 4-6-1, Shirokanedai Minato-ku, Tokyo 108, Japan

† Department of Anatomy and Brain Science, Kobe University School of Medicine, 7-5-1, Kusunoki-cho, Chuo-ku, Kobe 650, Japan

‡ Department of Molecular Biology, The Scripps Research Institute, MB10 10666 North Torrey Pines Road, La Jolla, California 92037, USA

Many biochemical, physiological and behavioural processes in organisms ranging from microorganisms to vertebrates exhibit circadian rhythms¹. In *Drosophila*, the gene *period* (*per*) is required for the circadian rhythms of locomotor activity and eclosion behaviour². Oscillation in the levels of *per* mRNA and Period (dPer) protein in the fly brain is thought to be responsible for the rhythmicity^{3,4}. However, no *per* homologues in animals other than insects have been identified. Here we identify the human and mouse genes (hPER and mPer, respectively) encoding PAS-domain (PAS, a dimerization domain present in Per, Amt and Sim)-containing polypeptides that are highly homologous to dPer. Besides this structural resemblance, mPer shows autonomous circadian oscillation in its expression in the suprachiasmatic

nucleus, which is the primary circadian pacemaker in the mammalian brain^{5,6}. *Clock*, a mammalian clock gene encoding a PAS-containing polypeptide^{7,8}, has now been cloned: it is likely that the *Per* homologues dimerize with other molecule(s) such as *Clock* through PAS–PAS interaction in the circadian clock system.

We developed a method of screening out short stretches of DNA sequences, called intra-module scanning–polymerase chain reaction (IMS–PCR) (see Methods) that allowed us to isolate mammalian homologues of the *Drosophila per* gene. Genetic and biochemical studies indicated the structural and functional significance of the PAS domain of the *Drosophila* protein dPer^{9,10}. We designed 18 combinations of primer pairs in the PAS-A and PAS-B repeats of dPer (Fig. 1a), and subjected human genomic DNA to PCR with the primers, and the products of PCR were then fractionated to yield DNA fragments of appropriate length (Fig. 1b, c; see Methods). The DNA bands of predicted lengths were then cloned and sequenced. Of 33 clones obtained from 12 bands (59–74 base pairs (bp)), a clone of 65 bp was specifically enriched sixfold to 21-fold in several nested PCRs with one primer pair (corresponding to the peptide sequences 5'-GYLPQD and 3'-FVHHEDI). The genomic DNA sequence encompassing the 65-bp fragment included an exon of 106 bp, which encodes 35 amino-acid residues within the entire 125 amino-acid PAS-B domain. A corresponding cDNA was isolated and designated as hPER cDNA. The mouse homologue (mPer) cDNA was cloned by using hPER cDNA as a probe. The hPER and mPer genes were mapped by FISH to chromosome 17p12-13.1 and 11B (data not shown), respectively, the syntenic loci between two animals.

The cDNA sequences of hPER and mPer contain open reading frames predicted to encode 1,290 and 1,291 amino-acid residues, respectively (Fig. 2b). Furthermore, the amino-acid sequences of hPer and mPer had 92% identity indicating that the genes hPER and mPer are conserved across these species (Fig. 2b). A homology search against a non-redundant amino-acid database using the BLAST program indicated that the two mammalian proteins show the highest homology score with dPer (type A)¹¹. Significant homology between the mammalian and *Drosophila* Per proteins is confined to five stretches (Fig. 2): (1) an amino-terminal homologous stretch (residues 44–131 of hPer and mPer, respectively); (2) PAS-A (residues 217–282 of both the Per homologues); (3) PAS-B (residues 338–456 of both the Per homologues) with adjacent

sequences downstream of PAS-B (residues 457–485 of both the Per homologues); (4) a short segment (residues 624–645 of both the Per homologues) corresponding to the region downstream of the site at which the *per*^S mutation (rendering a circadian period short) occurs (residue 589); and (5) a region homologous to the carboxy terminus of the Per-C region (residues 1,006–1,050 of hPer and 1,005–1,049 of mPer) followed by a serine–glycine (SG) repeat (residues 1,051–1,072 of hPer and 1,050–1,071 of mPer) and a homologous sequence further downstream of it (residues 1,073–1,108 of hPer and 1,072–1,107 of mPer). The extents of homology in these five stretches are 44%, 47%, 56%, 64% and 37%, respectively (Fig. 2a). Indeed, the PAS region (the stretches 2 and 3) in the Per homologue is the best relative of the correspondent region in dPer, although the other stretches also show strong homology. Five structural and functional domains have been characterized in dPer: (1) a nuclear localization signal (NLS) (residues 66–79)¹²; (2) the PAS domain (residues 233–490) required for interaction of dPer with the NLS of Tim¹⁰; (3) a cytoplasmic localization domain (CLD) mapped to the region downstream of the PAS-B repeat (residues 453–511)¹⁰; (4) the Per-C region (residues 524–685) that interacts with the PAS domain of its own polypeptide¹³; and (5) the threonine–glycine (TG) repeat (residues 694–748) and its immediate downstream region (residues 749–868) that regulate the rhythm of the species-specific courtship 'song' of *Drosophila*¹⁴. Thus the region homologous to NLS, PAS, CLD, two stretches in Per-C, and the TG repeat and its immediate C-terminal segment in each mammalian Per are aligned in exactly the same order as in dPer. In the C-terminal halves of the Per homologues, a short SG repeat is substituted for the TG repeat of dPer (Fig. 2b). This segment, surrounded by Per-C and the sequence homologous to the C-terminal vicinity of the TG repeat, is located approximately 350 residues downstream of the original position in dPer (Fig. 2a). These regions are also highly conserved in the human and mouse (Fig. 2b). Six Per segments (C1–C6) that are well conserved among *Drosophila* subspecies have been identified¹⁵ (Fig. 2a). As is the case for the Per homologue in the silkworm *Antheraea*, the portions of the mammalian Per homologous to dPer are restricted to the regions corresponding to C1–C3 of dPer¹⁶ (Fig. 2a). These findings lead us to conclude that hPER and mPer are the mammalian structural homologues of the *Drosophila* gene *per*.

Expression patterns of hPER and mPer were examined by north-

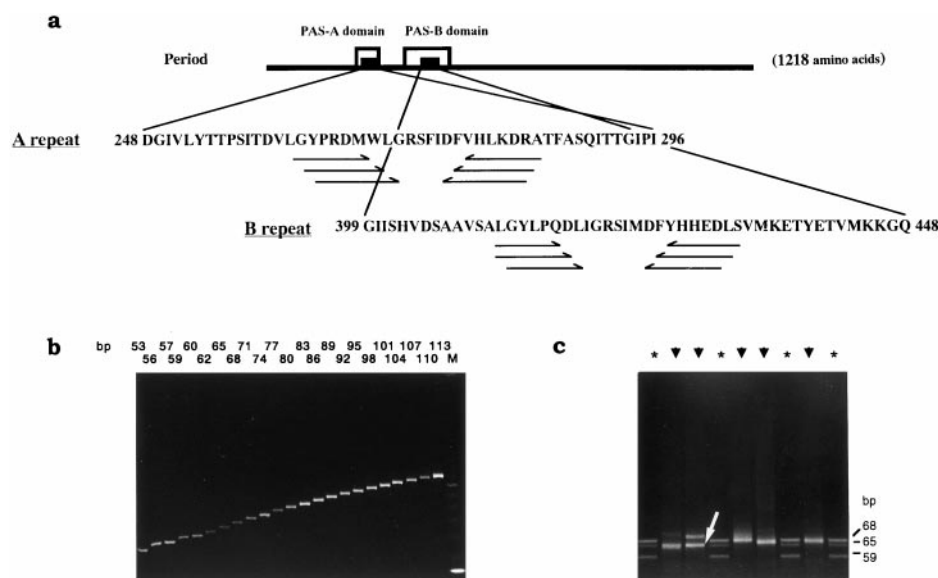


Figure 1 IMS–PCR (see Methods). **a**, Amino-acid sequences in the PAS repeats used to design the primer sequences are indicated by arrows. **b**, The 3-bp ladder marker was run on a 10% native PAGE gel under the discontinuous buffer condition²⁴. A 10-bp DNA ladder (BRL) was loaded into lane M. **c**, The PCR

products of IMS–PCR (lanes indicated by arrowheads) were run side by side with the 3-bp ladder marker (lanes indicated by asterisks) of 59, 65 and 68 bp. A 65-bp band encoding part of the PAS-B repeat of hPer is indicated by an arrow.

ern hybridization. A transcript of approximately 4.6 kilobases was detected in all adult human and mouse tissues examined, although the levels of *hPer* and *mPer* transcripts in comparison to those of *G3PDH* transcripts were different from each other (Fig. 3). This wide distribution of *hPer* and *mPer* expression was not surprising because *per* is expressed in tissues other than the brain in *Drosophila*^{17,18}.

The distribution of *mPer* mRNA in the mouse brain was examined by *in situ* hybridization. Weak signals of the *mPer* transcript were detected in most areas of the brain, including cortical and non-cortical structures. Stronger *mPer* mRNA signals were detected in the pyramidal cell layer of the piriform cortex, the periventricular part of the caudate-putamen, many of the thalamic nuclei, and the granular layer of the cerebellar cortex. The highest level of *mPer* expression in the brain was observed in the SCN at specific time points (Fig. 4a, b).

To explore the time dependence in the expression of *mPer* in the SCN, we entrained the mice by housing them in 12 h light : 12 h dark (LD). We quantified *mPer* mRNA in the SCN by *in situ* hybridization and by competitive RT-PCR, and both methods gave similar oscillation profiles in LD (Fig. 4a, c). The amount of *mPer* mRNA was highest in the light (at ZT4 to ZT8) (ZT, zeitgeber time; ZT0 = light on, ZT12 = light off), and lowest in the dark (at ZT16 to ZT20) (Fig. 4b). Furthermore, free-running fluctuations in the level of *mPer* mRNA occurred under conditions of constant darkness (DD) (Fig. 4a, c), being highest at CT4 to CT8 (CT, circadian time; CT0, subjective light on; CT12, subjective light off), and lowest at CT16 to CT20 (Fig. 4b). The maximum *mPer* mRNA level is approximately threefold higher than the minimum, under both LD and DD conditions (Fig. 4b). Thus the strong, autonomous and circadian expression of *mPer* mRNA in the SCN in constant darkness suggests that it acts as a circadian pacemaker, although the roles of the mammalian *per* homologues in the generation of circadian rhythms require further study using transgenic and

knock-out mice. The circadian cycling of *mPer* mRNA in the SCN resembles that of neural activity^{19–21}, peaking in the daytime and reaching a minimum at night. It is therefore conceivable that *mPer* regulates neuronal activity in the SCN.

Two mammalian circadian clock mutants have been established: *Clock* in the mouse²², and *tau* in the hamster²³. *Clock* seems to encode a transcription factor in the circadian clock^{5,6}, but *tau* has not yet been cloned. Analysis of the expression profiles of the *per* homologues in the SCN of the *Clock* and *tau* mutants, and the heterotopic dimerization of the *Per* homologues and *Clock*, should be of interest as both genetic and molecular interactions are expected to occur between the different elements in the same clock mechanism. The mammalian homologues of *Per* and *Clock* should lead to the identification of molecules that constitute the circadian clock, and may eventually lead to a better understanding of the regulatory mechanisms of circadian rhythms in higher organisms. □

Methods

IMS-PCR. In the human genome, short stretches of DNA sequences (modules) encoding short pieces of polypeptides (motifs) are dispersed over a long genomic distance. When a large number of 'intra-module scanning' (IMS) primers sufficient for covering an entire gene were used, the modules could be screened out by the polymerase chain reaction (PCR) at an equal frequency regardless of their expression levels. The sequences of the degenerate primer pairs of the PAS-A and PAS-B repeats were: GTGCTGGGCTACCCN(A/C)GNGA, CTGGGCTACCCCC(A/G)(A/G)GANATG, GGCTACCCCC(A/G)(A/G)GANATGTGG, CTGGGCT(A/T)CCTGCCNCA(A/G), CTGGGCT(A/T)CCTGCCNCA(A/G)GA, GGCTACCTGCC(C/T)CA(A/G)GAN(C/T), GCCCG(G/A)TCCTTCAG(G/A)TGNAC, TCCTCATG(A/G)TGCAC(A/G)(T/A)ANTC, ATGTCCTCATG(A/G)TG(C/G)AC(A/G)(A/T)A and GACAC(A/G)TCCTCATG(A/G)TG(A/G)TA, corresponding to the peptide sequences shown in Fig. 1a. When the corresponding amino-acid sequences were used for synthesizing PCR primers, the lengths of the PCR products should reflect the spatial characteristics of domain organization in the respective polypeptides, as homologous polypeptides share common features in comparable positions within the molecules. With the lengths of codons (3 bp each) and exons (average 100 bp each) in human genes taken into consideration, a 3-bp ladder marker (53–113 bp) was synthesized by PCR with a series of primers and pUC19 as a template (Fig. 1b). The marker was run side by side with the PCR products on a native PAGE (10%) gel in a discontinuous buffer system (Fig. 1c)²⁴. Each PCR mixture²⁵ contained 0.5 µg human genomic DNA. The mixture was incubated at 94 °C for 1 min, and subjected to 3 cycles of 94 °C for 30 s, 37 °C for 30 s and 72 °C for 30 s, followed by 25 cycles of 94 °C for 30 s, 45 °C for 30 s and 72 °C for 30 s.

Molecular cloning. A Marathon RACE kit, a PromoterFinder DNA walking kit, and poly(A)⁺ RNA extracted from human and mouse brains were purchased from Clontech. Northern hybridization was performed as described²⁶. A dye terminator cycle sequencing kit (ABI) was used for DNA sequencing with a 373A sequence (ABI).

Animals. *Balb/c* mice (male) 8 weeks old were used for all experiments. After 2 weeks of adaptation to the standard 12 h light : 12 h dark (LD) cycle, half the mice were transferred to an environment of constant darkness (DD) and kept there for 2 days, the rest being kept under LD.

In situ hybridization. Serial coronal sections (40 µm thick) of the mouse brain were made using a cryostat. *In situ* hybridization and quantification of mRNA have been described²⁷. The ³³P-labelled probes used for hybridization were the 5' sense and antisense strands of *mPer* cRNA (nucleotide positions 538–1752; data not shown). The radioactivity of each section on the BioMax film (Kodak) was analysed using a microcomputer interfaced to an image analysing system (MCID, Imaging Research) after conversion into the relative optical density using ¹⁴C-acrylic standards (Amersham). Data were normalized with respect to the difference between signal intensities in equal areas of the SCN and the corpus callosum. The intensities of the optical density of the sections from the rostralmost to the caudalmost of the SCN (10 sections per mouse) were then summed; the sum was considered a measure of the amount of *mPer* mRNA in this region.

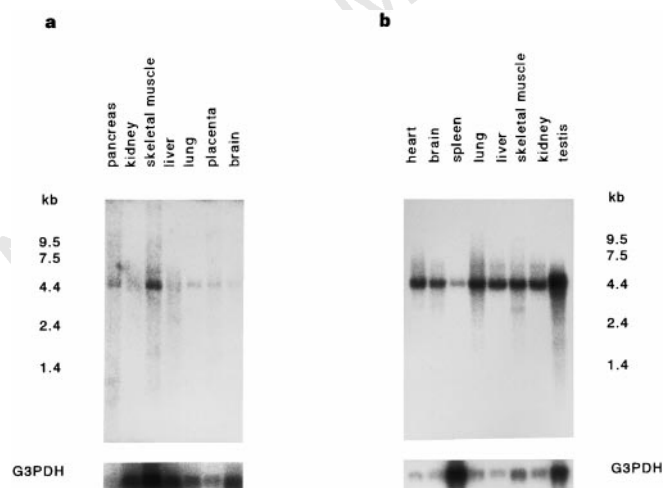


Figure 3 Northern blot analysis of *hPer* and *mPer*. The filters (a, human b, mouse) were purchased from Clontech. The filters were probed with *hPer* and *mPer*, respectively, and probed with *G3PDH* as a loading control.

Figure 2 Amino-acid sequence comparisons of the Period family. a, Homologous regions are boxed in different colours. C1–C6 represent the conserved regions of Period among *Drosophila* species¹⁵. b, The sequences corresponding to NLS, the PAS-A and PAS-B repeats, and CLD are underlined, and the TG (SG in *hPer* and *mPer*) repeat is boxed in red. Amino-acid identity between *hPer* and *mPer* is indicated by asterisks above the *hPer* sequence. Amino-acid identity and similarity between the mammalian and the *Drosophila* Period (*dPer*) are indicated by asterisks and circles below the *dPer* sequence, respectively.

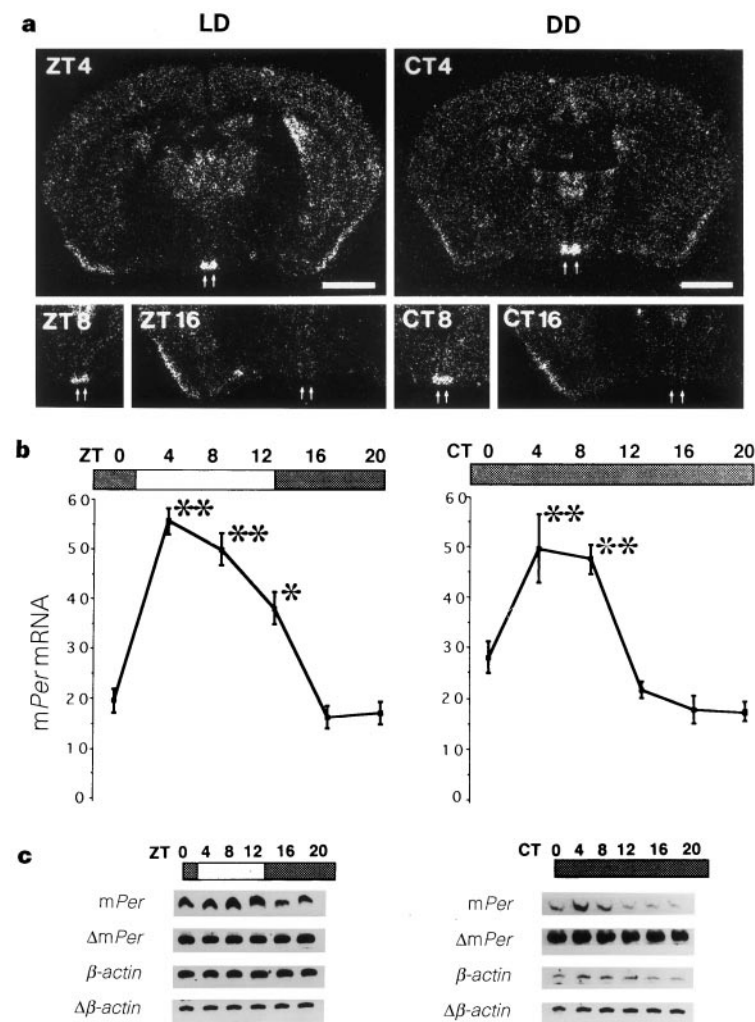


Figure 4 Autonomous cycling of mPer mRNA in the mouse SCN. **a**, *In situ* hybridization of mPer in the mouse brain under 12 h light: 12 h dark (LD, left) and total darkness (DD, right). The SCN is indicated by arrows. Scale bars, 2 mm. **b**, Quantification from *in situ* hybridization data. Values are means $\pm$ s.e.m. ($n = 5$). $**P < 0.01$; $*P < 0.05$, compared to the value at ZT16 or CT16 (ANOVA). **c**, Competitive RT-PCR analysis of mPer mRNA. Δ mPer, mPer competitor; $\Delta\beta$ -actin, β -actin competitor. In **b** and **c**, white bars indicate the duration of light, and grey bars the duration of dark.

Competitive RT-PCR analysis. Slices (0.5 mm thick) of mouse brain were obtained with Mouse Brain Matrix (Neuroscience, Tokyo). The SCN was punched out bilaterally from the frozen slices with a stereomicroscope with a microdissecting needle (gauge 600 μ m). Total RNA from the SCN ($n = 4$) was extracted using Trizol solution (BRL), treated with DNase I (Stratagene), and purified further with Trizol LS solution (BRL). A Superscript preamplification system (BRL) was used for reverse transcription of approximately 1 μ g of RNA, and mPer and β -actin cDNA were quantified by a competitive PCR method. PCR products were run on a native PAGE (5.5%) gel, and DNA in the appropriate bands was quantified with an FMBIOII fluorimager analyser (Hitachi) after staining with SYBR green (Molecular Probes). The competitor DNA fragments for mPer and β -actin were constructed by internal deletion of the respective cDNAs. The sizes of the PCR products of mPer, mPer competitor, β -actin, and β -actin competitor were 482, 246, 1,228 and 1,044 bp, respectively.

Received 8 May; accepted 29 July 1997.

- Edmunds, L. N. J. *Cellular and Molecular Basis of Biological Clocks* (Springer, New York, 1988).
- Konopka, R. J. & Benzer, S. Clock mutants of *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **68**, 2112–2116 (1971).
- Hardin, P. E., Hall, J. C. & Rosbash, M. Feedback of the *Drosophila period* gene product on circadian cycling of its messenger RNA levels. *Nature* **343**, 536–540 (1990).
- Zerr, D. M., Hall, J. C., Rosbash, M. & Siwicki, K. K. Circadian fluctuations of period protein immunoreactivity in the NS and the visual system of *Drosophila*. *J. Neurosci.* **10**, 2749–2762 (1990).
- Moore, R. Y. & Eichler, V. B. Loss of a circadian adrenal corticosterone rhythm following suprachiasmatic lesions in the rat. *Brain Res.* **42**, 201–216 (1972).
- Stephan, F. K. & Zucker, I. Circadian rhythms in drinking behavior and locomotor activity of rats are eliminated by hypothalamic lesions. *Proc. Natl Acad. Sci. USA* **69**, 1583–1586 (1972).
- King, D. P. *et al.* Positional cloning of the mouse circadian Clock gene. *Cell* **89**, 641–653 (1997).
- Antoch, M. P. *et al.* Functional identification of the mouse circadian Clock gene by transgenic BAC rescue. *Cell* **89**, 655–667 (1997).
- Bayliss, M. K., Bargiello, T. A., Jackson, F. R. & Young, M. W. Changes in abundance of structure of the *per* gene product can alter periodicity of the *Drosophila* clock. *Nature* **326**, 390–392 (1987).
- Saez, L. & Young, M. W. Regulation of nuclear entry of the *Drosophila* clock proteins Period and Timeless. *Neuron* **17**, 911–920 (1996).
- Citri, Y. *et al.* A family of unusually spliced biologically active transcripts encoded by a *Drosophila*

- clock gene. *Nature* **326**, 42–47 (1987).
- Vosshall, L. B., Price, J. L., Sehgal, A., Saez, L. & Young, M. W. Block in nuclear localization of period protein by a second clock mutation, timeless. *Science* **263**, 1606–1609 (1996).
- Huang, Z. J., Curtin, K. D. & Rosbash, M. PER protein interactions and temperature compensation of a circadian clock in *Drosophila*. *Science* **267**, 1169–1172 (1995).
- Wheeler, D. A. *et al.* Molecular transfer of a species-specific behavior from *Drosophila simulans* and *Drosophila melanogaster*. *Science* **251**, 1082–1085 (1991).
- Colot, H. V., Hall, J. C. & Rosbash, M. Interspecific comparison of the period gene of *Drosophila* reveals large blocks of non-conserved coding DNA. *EMBO J.* **7**, 3929–3937 (1988).
- Reppert, S. M., Tsai, T., Roca, A. L. & Sauman, I. Cloning of a structural and functional homolog of the circadian clock gene period from the giant silkworm *Antheraea pernyi*. *Neuron* **13**, 1167–1176 (1994).
- Liu, X., Lorenz, L., Yu, Q., Hall, J. C. & Rosbash, M. Spatial and temporal expression of the period gene in *Drosophila melanogaster*. *Genes Dev.* **2**, 228–238 (1988).
- Saez, L. & Young, M. W. In situ localization of the per clock protein during development of *Drosophila melanogaster*. *Mol. Cell. Biol.* **8**, 5378–5385 (1988).
- Inouye, S.-T. & Kawamura, H. Persistence of circadian rhythmicity in a hypothalamic 'island' containing the suprachiasmatic nucleus. *Proc. Natl Acad. Sci. USA* **76**, 5962–5966 (1979).
- Reppert, S. M., Tsai, T., Roca, A. L. & Sauman, I. Suprachiasmatic nucleus: use of 14 C-labeled deoxyglucose uptake as a functional marker. *Science* **197**, 1089–1092 (1977).
- Gillette, M. U. & Reppert, S. M. The hypothalamic suprachiasmatic nuclei: circadian patterns of vasopressin secretion and neuronal activity in vitro. *Brain Res. Bull.* **19**, 135–139 (1987).
- Vitaterna, M. H. *et al.* Mutagenesis and mapping of a mouse gene, *Clock*, essential for circadian behavior. *Science* **264**, 719–725 (1994).
- Ralph, M. R. & Menaker, M. A mutation of the circadian system in golden hamsters. *Science* **241**, 1225–1227 (1988).
- Ito, T., Hohjoh, H. & Sakaki, Y. Pulse-field polyacrylamide gel electrophoresis: basic phenomena and applications. *Electrophoresis* **14**, 278–282 (1993).
- Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning* (Cold Spring Harbor Laboratory Press, NY, 1989).
- Church, G. M. & Gilbert, W. Genomic sequencing. *Proc. Natl Acad. Sci. USA* **81**, 1991–1995 (1984).
- Ban, Y., Shigeyoshi, Y. & Okamura, H. Development of circadian VIP rhythm in the rat suprachiasmatic nucleus. *J. Neurosci.* **17**, 3920–3931 (1997).

Acknowledgements. We thank J. Inazawa and M. Hirai for mapping hPER and mPer, respectively; Y. Maebayashi for technical assistance; S.-T. Inouye and M. Hattori for discussion; and D. Yamamoto, R. Ueda, T. Ito and K. Ui-Tei for comments and discussions on the manuscript. This work was partly supported by research grants from the Japanese Ministry of Education, Science, Sports and Culture, and the Japanese Ministry of Health and Welfare.

Correspondence and requests for materials should be addressed to H.T. (e-mail: tei@ims.u-tokyo.ac.jp). The sequences for hPER and mPer are deposited in the DDBJ database (accession nos AB002107 and AB002108, respectively).

Head induction by simultaneous repression of Bmp and Wnt signalling in *Xenopus*

Andrei Glinka, Wei Wu, Darya Onichtchouk, Claudia Blumenstock & Christof Niehrs

Division of Molecular Embryology, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany

The Spemann organizer of the amphibian embryo can be subdivided into two discrete activities, namely trunk organizer and head organizer¹. Several factors secreted from the organizer that are involved in trunk organization are thought to act by repressing Bmp signalling²⁻⁴. With the exception of the secreted factor *cerberus*⁵, little is known about head-organizer inducers. Here we show that co-expression of a dominant-negative Bmp receptor with inhibitors of the Wnt-signalling pathway in *Xenopus* leads to the induction of complete secondary axes, including a head. This induction does not require expression of the *siamois* marker of Nieuwkoop centre signalling, suggesting that cells are directly

shifting to head-organizer fate. Furthermore, we find that *cerberus* is a potent inhibitor of Wnt signalling. Our results indicate that head-organizer activity results from the simultaneous repression of Bmp and Wnt signalling and they suggest a mechanism for region-specific induction by the organizer.

In *Xenopus*, the Wnt-signalling pathway is involved in at least two different processes in mesoderm formation. First, maternal activation of the pathway is required for the initiation of axis formation, by creating a Nieuwkoop centre. This activation can occur in the presence of certain Wnt inhibitors and thus may occur by direct activation of a downstream component of the Wnt pathway⁶⁻⁹. Second, the Wnt pathway is zygotically activated after midblastula transition (MBT) through *Xwnt-8*, which is expressed in and required for ventrolateral mesoderm formation¹⁰. Ubiquitous *Xwnt-8* expression after MBT interferes with head development¹¹, and expression of a dominant-negative form of *Xwnt-8* (dnXwnt-8) leads to the formation of embryos with a shortened primary axis and enlarged heads⁹ (Fig. 1a, b). This suggests that repression of Wnt signalling may promote head formation. However, ectopic heads and secondary embryonic axes are never observed when embryos are injected with dnXwnt-8 on the ventral side, raising the possibility that the ventral side contains inhibitors other than Wnt signals, which prevent ectopic head formation.

Bone morphogenetic proteins (BMPs) such as *Bmp-4* and *Bmp-7*, which are coexpressed with *Xwnt-8* in the ventral and lateral regions of the *Xenopus* gastrula, antagonize trunk organizer activity^{2,3,4}

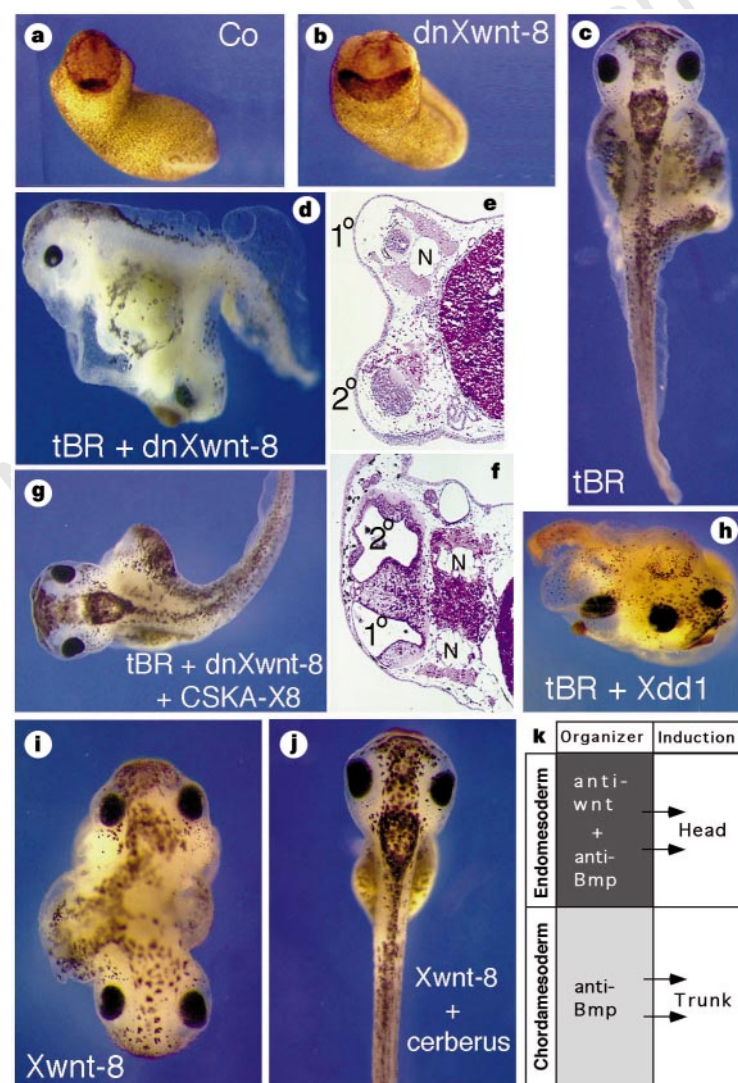


Figure 1 Head induction by simultaneous inhibition of Bmp and Wnt signalling. Four-cell embryos were uninjected (**a**) (control), radially injected into all blastomeres (**b**) or injected in two ventral blastomeres with the RNA or DNA indicated. The final dose of mRNAs per blastomere were 1.2 ng (**b**) or 0.3 ng (**d, f, g**) dnXwnt-8, 0.25 ng tBR, 0.5 ng Xdd1, 0.0015 ng Xwnt-8 mRNA, 0.025 ng *cerberus* or 0.1 ng CSKA-X8 DNA, as indicated. Embryonic phenotypes resulting from the injections are shown in **b, c, d, g-j**. Transverse histological sections²⁹ of secondary embryonic axes induced by tBR and tBR (0.25 ng) + dnXwnt-8 (0.1 ng) are shown in **e** and **f**, respectively. Primary and secondary axes are indicated. Note that embryos injected with tBR/dnXwnt-8 show notochord (N) in the secondary axis (80% by MZ15 immune-wholemount staining; $n = 15$). **k**, Model for region-specific anteroposterior induction by the Spemann organizer. Bmp and Wnt signals are required for trunk-tail formation. Trunk induction requires inhibition of Bmp signalling; head induction requires dual inhibition of Wnt and Bmp signalling. The organizer produces factors that inhibit both types of signals (anti-Wnt, anti-Bmp). Regional specificity of induction results from differential expression of Wnt-inhibiting signals such as *frzb*^{7,8} and *cerberus*⁵ in endomesoderm and Bmp-inhibiting signals such as *noggin*¹⁴, *chordin*²⁵ and *folistatin*²⁶ in endomesoderm and chordamesoderm. The model is highly simplifying and omits many other factors.

Expression of the dominant-negative Bmp-2/4 receptor (tBR) overcomes this inhibition and induces secondary embryonic axes that lack a complete head and notochord^{12,13} (Fig. 1c, e, and Table 1, experiment I-1). To test whether head formation requires simultaneous repression of Bmp and Wnt signalling, tBR¹³ was co-injected with dnXwnt-8 into the ventral side of 4-cell embryos. This caused the formation at high frequency of complete secondary embryonic axes with notochords, and heads with eyes and cement glands (Fig. 1d, f, and Table 1, series I). Induced heads typically contained only one eye. This could be due to incomplete inhibition of Wnt/Bmp signalling, although neither higher doses of dnXwnt-8, nor tBR in combination, resulted in the formation of secondary axes with two eyes (Table 1, series I). To determine whether repression of Bmp and Wnt signalling is also effective after MBT, we injected pCS-noggin, which encodes a secreted inhibitor of Bmp signalling¹⁴, together with pCS-dnXwnt-8 (ref. 9). In these plasmids, the expression of both genes is under the control of the cytomegalovirus (CMV) promoter, which becomes activated after MBT. This injection also induced complete axes, albeit at lower frequency (Table 1, experiment II-3), indicating that repression of zygotic Wnt and Bmp signalling are responsible for the observed effects.

Complete axis induction by dnXwnt-8 is due to specific inhibition of Wnt signalling. First, it can be rescued by co-injecting a plasmid (CSKA-X8) encoding wild-type *Xwnt-8* under the control of the cytoskeletal actin promoter, which is activated after MBT¹¹

(Fig. 1g, and Table 1, series III). Second, co-injection of tBR with dominant-negative *dishevelled* (Xdd1) messenger RNA⁶, an inhibitor of a downstream Wnt-signalling component, also leads to the formation of complete secondary heads with single eyes (Fig. 1h, and Table 1, series III). We conclude that simultaneous repression of Bmp and Wnt signalling confers head-organizer activity in *Xenopus*.

Head organizer is induced by the Nieuwkoop centre via downstream components of the maternal Wnt signalling cascade¹⁰. *Siamois* (*sia*) is a target of this dorsalization pathway and triggers organizer activity¹⁵. To determine whether head formation by repression of Bmp and Wnt signalling occurs via a step involving *sia* expression, we assayed the induction of *sia*. Figure 2a shows that ventral marginal zone (VMZ) explants microinjected with mRNA encoding *Xwnt-8*, an artificial inducer of the Nieuwkoop centre^{16,17}, express *sia*. In addition, these explants express the dorsoanterior markers *otx2* (refs 18, 19), *gsc* (ref. 20) and *Xlim-1* (ref. 21). In contrast, VMZs co-injected with tBR/dnXwnt-8 mRNAs do not express *sia* but do express *otx2*, *Xlim-1* and *gsc*. The mesodermal marker *Xbra*²² is downregulated by injection of all mRNAs. We found that dnXwnt-8 alone is able to reduce the anterior neural markers *opsin*, *otx2*, *XAG1* and *engrailed* in animal cap explants (Fig. 2b). This induction occurs in the absence of both mesodermal gene expression (*Xbra*, actin) and overt neural histodifferentiation (data not shown), and suggests that Wnt signalling normally prevents neural gene expression in the animal cap²³.

These results indicate that complete axis induction by repression of Bmp and Wnt signalling occurs by directly conferring head-organizer activity without involving a pathway requiring *sia* expression. This suggests that one function of head inducers may be to inhibit Wnt signalling. To test this for the head inducer *cerberus*, we investigated whether secondary axis formation by *Xwnt-8* mRNA injection could be rescued by co-injection of *cerberus* mRNA, an assay used previously to test for Wnt-antagonizing activity⁶⁻⁹. We found that even doses of *cerberus* that by themselves had no phenotypic effect dramatically abolished secondary-axis formation by *Xwnt-8* (Fig. 1i, j, and Table 1, series IV). At ten-times-higher dose, *Xwnt-8* can titrate out *cerberus* and reinduce secondary axes (Table 1 IV-6), suggesting that *cerberus* and *Xwnt-8* can antagonize each other directly or indirectly. *Cerberus* is a head inducer and is shown here to be a potent inhibitor of Wnt signalling. However, because head-organizer activity also requires inhibition of Bmp signalling, *cerberus* should do more than inhibit Wnt signals. As *cerberus* can override the antineuralizing effect of endogenous *Bmp-4* in animal caps⁵ the gene may inhibit both Wnt and Bmp signalling pathways. Consistent with this, we find that *cerberus* RNA injection downregulates expression of the *Bmp-4* target gene *Xvent-1* (ref. 24) (data not shown), but there are significant differences between *cerberus* and tBR/dnXwnt-8 action. First, *cerberus* induces only heads, whereas tBR/dnXwnt-8 induce a secondary embryonic axis with notochord and head. In fact, *cerberus* inhibits axis formation (ref. 5 and data not shown). Second, *cerberus* does not enhance head formation after co-injection with tBR/dnXwnt-8, as would be expected if it acts in the same way, but inhibits it (Table 1, IV-7). Thus, the mechanisms of action of tBR/dnXwnt-8 and *cerberus* only partially overlap: other activities of *cerberus* or different specificities for inhibitors of Wnt and TGF- β signalling may account for these differences.

Our evidence shows that the functional division between trunk- and head-organizer activity can be attributed to inhibition of Bmp signalling (trunk only) and dual repression by Bmp as well as Wnt signalling (trunk plus head). These findings are supported by the recent identification of *frzb*, which can antagonize Wnt protein by direct binding^{7,8}. *Frzb* is a candidate factor for a Wnt-inhibiting head inducer because it is predominantly expressed in the involuting endomesoderm, the part of the Spemann organizer that induces head structures¹ which also expresses *cerberus*⁵. *Frzb* is co-expressed

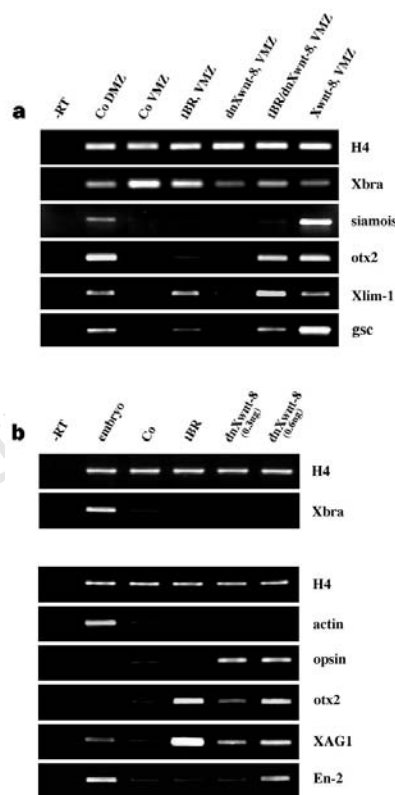


Figure 2 Induction of head organizer by inhibition of Bmp and Wnt signalling is direct. **a**, VMZ assay: 4-cell embryos were microinjected into all blastomeres with 0.25 ng tBR, 0.3 ng dnXwnt-8 or a combination of both, as indicated. Xwnt-8 mRNA was at 1 pg/blastomere. Dorsal (DMZ) or ventral marginal zone (VMZ) fragments were cut from early gastrulae and analysed at stage 10.5 by RT-PCR. Note the induction of *siamois* by Xwnt-8 but not by tBR/dnXwnt-8. **b**, Animal cap assay: 4-cell embryos were microinjected into the animal region of all blastomeres with 0.25 ng tBR, or dnXwnt-8 as indicated. Animal cap fragments were cut from late blastulae, cultivated until stage 10.5 (top) or stage 32 equivalent (bottom) and analysed by RT-PCR for expression of marker genes indicated. H4, histone H4; En-2, engrailed-2; -RT, minus reverse transcription control sample.

Table 1 Head induction by simultaneous repression of Bmp and Wnt signalling

Expt.	Injected (dose)	n	Incomplete s.a. (%)	Secondary head (%)
I-1	tBR (0.25)	105	35	0
I-2	dnXwnt-8 (0.15–1.2)	113	0	0
I-3	dnXwnt-8 (0.1) + tBR (0.25)	118	12	45
I-4	dnXwnt-8 (0.3) + tBR (0.25)	85	17	40
I-5	dnXwnt-8 (0.6) + tBR (0.25)	26	11	42
I-6	dnXwnt-8 (0.3) + tBR (1)	51	0	25
I-7	dnXwnt-8 (0.6) + tBR (1)	38	26	39
II-1	CS-noggin (0.005)	52	46	1
II-2	CS-dnXwnt-8 (0.05)	35	3	0
II-3	CS-noggin (0.005) + CS-dnXwnt-8 (0.05)	35	14	26
III-1	dnXwnt-8 (0.1) + tBR (0.25) + CSKA-X8 (0.1)	137	36	0.7
III-2	Xdd1 (0.5)	50	2	0
III-3	Xwnt-8 (0.0015)	25	12	60
III-4	Xwnt-8 (0.0015) + Xdd1 (0.5)	39	0	0
III-5	Xdd1 (0.5) + tBR (0.25)	62	3	53
IV-1	Xwnt-8 (0.0015)	50	0	66
IV-2	Xwnt-8 (0.015)	47	0	40
IV-3	cerberus (0.025)	108	0	0
IV-4	cerberus (0.2)	57	0	72
IV-5	Xwnt-8 (0.0015) + cerberus (0.025)	57	2	0
IV-6	Xwnt-8 (0.015) + cerberus (0.025)	45	0	40
IV-7	dnXwnt-8 (0.3) + tBR (0.25) + cerberus (0.025)	45*	4	11

Four-cell embryos were microinjected into two ventral blastomeres with the mRNAs or cDNAs (pCS-noggin, pCS-dnXwnt-8, pCSKA-X8) indicated (doses are in ng per blastomere). Embryos were scored for the formation of incomplete secondary embryonic axes (s.a.) or secondary axes containing secondary heads with eyes and cement glands. n, Total number of embryos analysed. A minimum of two independent experiments were typically carried out for every experiment. Series I, induction of head organizer by inhibition of Bmp and Wnt signalling; series II, head induction occurs by removal of zygotic signals; series III, specificity of Wnt-signalling repression; series IV, Cerberus is a potent inhibitor of Wnt signalling.

* 24% of embryos contained ectopic patches of cement gland.

in endomesoderm with the Bmp-signalling inhibitors *noggin*¹⁴, *chordin*²⁵ and *folistatin*²⁶, and induces complete secondary axes in combination with tBR (data not shown). Wnt and Bmp signalling are required for trunk–tail patterning and the organizer produces secreted factors that inhibit both. As shown here, in the complete absence of trunk–tail signalling, cells shift to head-organizer fate. In the absence of only Bmp signalling, cells shift to trunk organizer. This suggests a model for region-specific anteroposterior induction by the Spemann organizer (Fig. 1k). Our discovery of the dual repression of Wnt/Bmp signalling in the head organizer should aid in the identification of other head inducers. □

Methods

Embryos and explants. *In vitro* fertilization, embryo culture, staging, microinjection and culture of explants were carried out as described²⁷.

Constructs and microinjection experiments. Full-length *cerberus* was amplified by using the polymerase chain reaction (PCR) from a *Xenopus* neurula cDNA library²⁷ and cloned into pCS2²⁸ as described⁵. pCS-noggin was constructed by cloning an *EcoRI* (blunt)–*HindIII* (blunt) fragment of *noggin* Δ5¹⁴ into *XhoI* (blunt) of pCS2. Plasmids were linearized and capped mRNA was transcribed as follows: dnXwnt-8 (ref. 9) (*Asp*718, Sp6), tBR¹³ (*EcoRI*, Sp6), Xdd1 (ref. 6) (*NorI*, Sp6), Xwnt-8 (ref. 16) (*Bam*H1, Sp6), *cerberus* (*Asp*718, Sp6).

RT-PCR. PCR assays with reverse transcription (RT-PCR) were carried out in the exponential phase of amplification as described²⁴. PCR primers used were *actin*²⁹, *engrailed* (*En*-2)²⁶, *gsc*²⁷, histone H4 (ref. 27), *opsin*²⁶, *otx*-2 (ref. 18, *siamois* (f: AAACCACTGATTTCAGGCAGAGG; r: GTAGGGCTGTGTATTT-GAAGGG), *XAG1* (ref. 30), *Xbra*²⁷ and *Xlim*-1 (f: ACTGACTTCTTCAGGA-GATTTGG; r: GTTCTCGCCTGTTGAGAGC).

Received 2 June; accepted 10 July 1997.

- Gerhart, J., Doniach, T. & Steward, R. in *Gastrulation* (eds Keller, R., Clark, W. H. & Griffin, F.) 57–77 (Plenum, New York, 1991).
- De Robertis, E. M. & Sasai, Y. A common plan for dorsoventral patterning in *Bilateria*. *Nature* **380**, 37–40 (1996).
- Lemaire, P. & Kodjabachian, L. The vertebrate organizer: structure and molecules. *Trends Genet.* **12**, 525–531 (1996).
- Hemmati-Brivanlou, A. & Melton, D. Vertebrate embryonic cells will become nerve cells unless told otherwise. *Cell* **88**, 13–17 (1997).
- Bouwmeester, T., Kim, S.-H., Sasai, Y., Lu, B. & De Robertis, E. M. Cerberus is a head-inducing secreted factor expressed in the anterior endoderm of Spemann's organizer. *Nature* **382**, 595–601 (1996).
- Sokol, S. Y. Analysis of dishevelled signalling pathways during *Xenopus* development. *Curr. Biol.* **6**, 1456–1467 (1996).
- Wang, S., Krinks, M., Lin, K., Luyten, F. P. & Moos, M. Frzb, a secreted protein expressed in the

- Spemann organizer, binds and inhibits Wnt-8. *Cell* **88**, 757–766 (1997).
- Leyns, L., Bouwmeester, T., Kim, S.-H., Piccolo, S. & De Robertis, E. M. Frzb-1 is a secreted antagonist of wnt-signals expressed in the Spemann organizer. *Cell* **88**, 747–756 (1997).
- Hoppler, S., Brown, J. D. & Moon, R. T. Expression of a dominant-negative wnt blocks induction of MyoD in *Xenopus* embryos. *Genes Dev.* **10**, 2805–2817 (1996).
- Miller, J. & Moon, R. T. Signal transduction through beta-catenin and specification of cell fate during embryogenesis. *Genes Dev.* **10**, 2527–2539 (1996).
- Christian, J. L. & Moon, R. T. Interactions between Xwnt-8 and Spemann organizer signaling pathways generate dorsoventral pattern in the embryonic mesoderm of *Xenopus*. *Genes Dev.* **7**, 13–28 (1993).
- Graff, J. M., Thies, R. S., Song, J. J., Celeste, A. J. & Melton, D. A. Studies with a *Xenopus* BMP receptor suggest that ventral mesoderm-inducing signals override dorsal signals *in vivo*. *Cell* **79**, 169–179 (1994).
- Suzuki, A. et al. A truncated bone morphogenetic protein receptor affects dorsal-ventral patterning in the early *Xenopus* embryo. *Proc. Natl. Acad. Sci. USA* **91**, 10255–10259 (1994).
- Smith, W. C. & Harland, R. M. Expression cloning of *noggin*, a new dorsalizing factor localized to the Spemann organizer in *Xenopus* embryos. *Cell* **70**, 829–840 (1992).
- Carnac, G., Kodjabachian, L., Gurdon, J. B. & Lemaire, P. The homeobox gene *siamois* is a target of the wnt dorsalization pathway and triggers organizer activity in the absence of mesoderm. *Development* **122**, 3055–3065 (1996).
- Sokol, S., Christian, J. L., Moon, R. T. & Melton, D. A. Injected Wnt RNA induces a complete body axis in *Xenopus* embryos. *Cell* **67**, 741–752 (1991).
- Smith, W. C. & Harland, R. M. Injected Xwnt-8 RNA acts early in *Xenopus* embryos to promote formation of a vegetal dorsalizing center. *Cell* **67**, 753–765 (1991).
- Blitz, I. L. & Cho, K. W. Y. Anterior neuroectoderm is progressively induced during gastrulation: the role of the *Xenopus* homeobox gene *orthodenticle*. *Development* **121**, 993–1004 (1995).
- Pannese, M. et al. The *Xenopus* homologue of *Otx2* is a maternal homeobox gene that demarcates and specifies anterior body regions. *Development* **121**, 707–720 (1995).
- Cho, K. W., Blumberg, B., Steinbeisser, H. & De Robertis, E. M. Molecular nature of Spemann's organizer: the role of the *Xenopus* homeobox gene *gooseoid*. *Cell* **67**, 1111–1120 (1991).
- Taira, M., Jamrich, M., Good, P. J. & Dawid, I. B. The LIM domain-containing homeobox gene *Xlim-1* is expressed specifically in the organizer region of *Xenopus* gastrula embryos. *Genes Dev.* **6**, 356–366 (1992).
- Smith, J. C., Price, B. M., Green, J. B., Weigel, D. & Herrmann, B. G. Expression of a *Xenopus* homolog of Brachyury (T) is an immediate-early response to mesoderm induction. *Cell* **67**, 79–87 (1991).
- Itoh, K., Tang, T. L., Neel, B. G. & Sokol, S. Y. Specific modulation of ectodermal cell fates in *Xenopus* embryos by glycogen synthase kinase. *Development* **121**, 3979–3988 (1995).
- Gawantka, V., Delius, H., Hirschfeld, K., Blumenstock, C. & Niehrs, C. Antagonizing the Spemann organizer: role of the homeobox gene *Xvent-1*. *EMBO J.* **14**, 6268–6279 (1995).
- Sasai, Y. et al. *Xenopus* chordin: a novel dorsalizing factor activated by organizer-specific homeobox genes. *Cell* **79**, 779–790 (1994).
- Hemmati-Brivanlou, A., Kelly, O. G. & Melton, D. A. Follistatin, an antagonist of activin, is expressed in the Spemann organizer and displays direct neuralizing activity. *Cell* **77**, 283–295 (1994).
- Glinka, A., Delius, H., Blumenstock, C. & Niehrs, C. Combinatorial signalling by Xwnt-11 and Xnr3 in the organizer epithelium. *Mech. Dev.* **60**, 221–231 (1996).
- Rupp, R. A., Snider, L. & Weintraub, H. *Xenopus* embryos regulate the nuclear localization of XMyoD. *Genes Dev.* **8**, 1311–1323 (1994).
- Niehrs, C., Steinbeisser, H. & De Robertis, E. M. Mesodermal patterning by a gradient of the vertebrate homeobox gene *gooseoid*. *Science* **263**, 817–820 (1994).
- Gammil, L. S. & Sive, H. Identification of *otx2* target genes and restrictions in ectodermal competence during cement gland formation. *Development* **124**, 471–481 (1996).

Acknowledgements. We thank R. Harland, R. Moon, S. Sokol and M. Ueno for constructs, and P. Monaghan for critically reading the manuscript. This work was supported by the DFG.

Correspondence and requests for materials should be addressed to C.N. (e-mail: niehrs@dkfz-heidelberg.de).

Protein memory through altered folding mediated by intramolecular chaperones

U. P. Shinde, J. J. Liu & M. Inouye

Department of Biochemistry, Robert Wood Johnson Medical School-UMDNJ, 675 Hoes Lane, Piscataway, New Jersey 08854, USA

The 77-residue propeptide of subtilisin acts as an intramolecular chaperone that organizes the correct folding of its own protease domain. Similar folding mechanisms are used by several prokaryotic and eukaryotic proteins, including prohormone-convertases. Here we show that the intramolecular chaperone of subtilisin facilitates folding by acting as a template for its protease domain, although it does not form part of that domain. Subtilisin E folded by an intramolecular chaperone with an Ile(−48)-to-Val mutation acquires an 'altered' enzymatically active conformation that differs from wild-type subtilisin E. Although both the altered and wild-type subtilisins have identical amino-acid sequences, as determined by amino-terminal sequencing and mass spectrometry, they bind their cognate intramolecular chaperones with 4.5-fold greater affinity than non-cognate intramolecular chaperones, when added *in trans*. The two subtilisins also have different secondary structures, thermostability and substrate specificities. Our results indicate that an identical polypeptide can fold into an altered conformation through a mutated intramolecular chaperone and maintains memory of the folding process. Such a phenomenon, which we term 'protein memory', may be important in investigations of protein folding.

The amino-acid sequence encodes all the information necessary for a protein to adopt its three-dimensional structure¹. Protein folding describes the transition of a non-functional polypeptide into a functional protein molecule². Some proteins require an intramolecular chaperone (IMC) to act as a single turn-over

catalyst to mediate correct folding^{3–7}. Upon folding, the IMC is removed through an autolytic³ or exogenous proteolytic activity. Subtilisin^{3,7–10}, α -lytic protease¹¹, aqualysin¹² and carboxypeptidase Y¹³ constitute some well-studied examples of such proteins. Subtilisin undergoes conformational changes upon autoprocessing of its IMC domain¹⁴. At low concentrations and in the absence of their IMC domains, subtilisin and α -lytic protease acquire molten-globule-like intermediate states^{15–17}. Addition of IMC domains *in trans* promotes these intermediates to surmount the activation energy barrier between the intermediate and the native states^{15,17}, suggesting that IMCs help overcome kinetic barriers during the late stages of folding^{15,16,18}. They also function as temporary inhibitors of enzymatic activity^{19,20}.

The *in vitro* maturation pathway of pro-subtilisin E involves folding, autoprocessing and degradation of its 77-residue IMC domain, resulting in a 275-residue protease domain. An I−48V mutation within the IMC has been isolated²¹. Mutant and wild-type pro-subtilisins and IMC domains were cloned, expressed and purified in *Escherichia coli* using vectors pH1215T (ref. 21) and pET11a (ref. 22). Proteins were refolded and then analysed using SDS-PAGE (polyacrylamide gel electrophoresis). Wild-type and mutant precursors autoprocess and degrade their IMC domains to generate the enzymatically active proteins wild-type subtilisin (Sub^{WT}) and the 'altered' subtilisin (Sub^{PC}) (Fig. 1a). Both Sub^{WT} and Sub^{PC} display similar retention times on an analytical gel-filtration column (Fig. 1b), suggesting their corresponding Stokes radii are similar. Moreover, the proteins do not contain partly or completely unfolded polypeptides. Mass spectroscopy shows that the relative molecular mass (M_r) of both Sub^{PC} ($27,719 \pm 10$) and Sub^{WT} ($27,728 \pm 10$) are within range of the theoretical M_r of subtilisin E (27,720), expressed using the vector pH1215T (ref. 21). Amino-terminal sequencing shows that both proteins have identical amino termini. Circular dichroism spectra (Fig. 1c) indicate that Sub^{WT} and Sub^{PC} display similar but not identical secondary structures and represent homogenous but structurally distinct folded protein populations.

The catalytic properties of Sub^{WT} and Sub^{PC} were estimated using *N*-succinyl-AAPF-*p*-nitroanilide as a substrate^{3,24}. Sub^{WT} and Sub^{PC}

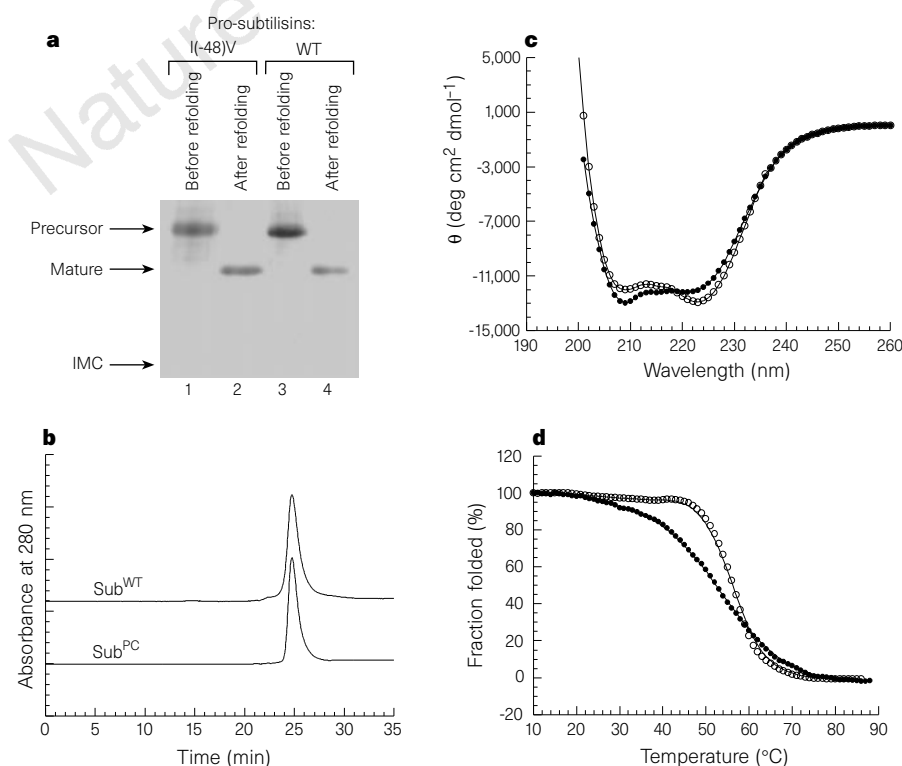


Figure 1 a, Maturation of wild-type (WT) and Ile(−48)Val mutant pro-subtilisin. The refolded proteins were electrophoresed on a 17.5% acrylamide gel. Lanes 1 and 2 represent the Ile(−48)Val mutant, and lanes 3 and 4 represent wild-type pro-subtilisins before and after refolding. **b**, Analytical gel-filtration chromatography of Sub^{WT} and Sub^{PC} after refolding. Sub^{PC} (lane 2 in **a**) and Sub^{WT} (lane 4 in **a**) were applied on a TSK2000 column and eluted using 10 mM phosphate buffer, pH 7.0, containing 0.5 M ammonium sulphate, 1 mM CaCl₂, with a flow rate of 0.5 ml min^{−1}. **c**, Circular-dichroism spectra of Sub^{WT} (open circles) and Sub^{PC} (filled circles). Temperature was maintained at 25 °C. **d**, Thermal unfolding of Sub^{WT} and Sub^{PC}. The fraction unfolded was monitored using circular dichroism and recording the changes in negative ellipticity at 222 nm as a function of temperature. Unfolding of Sub^{PC} (filled circles) is less cooperative than Sub^{WT} (open circles).

Table 1 Properties of Sub^{WT} and Sub^{PC}

Property	Sub ^{WT}	Sub ^{PC}
K_M for succinyl-AAPF- <i>p</i> -nitroanilide ($\times 10^{-3}$ M)	2.040 ± 0.015	0.675 ± 0.005
K_{cat} for succinyl-AAPF- <i>p</i> -nitroanilide (s^{-1})	$22,000 \pm 0.15$	$11,800 \pm 0.080$
K_{cat}/K_M ($s^{-1} \times 10^3$ M)	10.8	17.5
Apparent K_i for IMC ^{WT} ($\times 10^{-9}$ M)	153	694
Apparent K_i for IMC ^{PC} ($\times 10^{-9}$ M)	846*	184*
Activity for succinyl-AAPF- <i>p</i> -nitroanilide (%)	100	54
Activity for succinyl-AAA- <i>p</i> -nitroanilide (%)	100	71
Activity for succinyl-AAL- <i>p</i> -nitroanilide (%)	100	88

* Calculated from data for 600 s because after this time rapid degradation of the IMC occurs.

display a K_M of 2.040 ± 0.015 and 0.675 ± 0.005 mM, respectively (Table 1). The K_{cat}/K_M ratio for Sub^{PC} is 62% greater than that for Sub^{WT}. Interactions between subtilisin and the IMC follow slow-binding inhibition kinetics^{22,24,25}. Estimated values of the inhibition constant K_i of IMC^{WT} and IMC^{PC} for Sub^{WT} and Sub^{PC} show that IMC^{PC} binds Sub^{WT} with an apparent K_i that is 4.6-fold higher than its apparent K_i for Sub^{PC}, whereas IMC^{WT} binds Sub^{PC} 4.5-fold weaker than with Sub^{WT} (Table 1). If Sub^{WT} and Sub^{PC} are denatured and refolded using IMC^{WT} *in trans*, both display similar kinetics towards *N*-succinyl-AAPF-*p*-nitroanilide. Incubating folded Sub^{PC} with IMC^{WT} did not change the catalytic properties of Sub^{PC} to Sub^{WT}.

The thermostabilities of Sub^{WT} and Sub^{PC} were determined by monitoring changes in negative ellipticity at 222 nm (Fig. 1d). Sub^{PC} was found to be relatively unstable (melting temperature, T_m , 52 °C) and less cooperative than Sub^{WT} (T_m , 58.5 °C). The X-ray structure of the IMC^{WT}-subtilisin complex²³ leads us to speculate on domains that may be affected. There are important interactions between the IMC domain and subtilisin between residues 100 and 144. Ile(-48) in the IMC is buried inside a hydrophobic core constituted by four β -sheets and two α -helices. The I-48V substitution may create a cavity within the hydrophobic core which perturbs the final structure of the protease domain. Structural changes resulting from mutations of surface residues

in lysozyme are localized near the mutation, whereas those involving buried residues may be transmitted to other parts of the protein²⁶. Similar structural alterations may affect surface charge distribution in subtilisin, which can alter its enzymatic properties²⁷. The relative activities of Sub^{WT} and Sub^{PC} towards *N*-succinyl-AAPF-*p*-nitroanilide, *N*-succinyl-AAA-*p*-nitroanilide and *N*-succinyl-AAL-*p*-nitroanilide are different (Table 1).

Protease domains degrade IMCs upon completing protein folding²². To further characterize Sub^{WT} and Sub^{PC}, we analysed their interactions with IMC^{WT} and IMC^{PC}. IMC^{WT} acquires significant secondary structure only after interacting with subtilisin^{7,15,23}, and slow-binding inhibition kinetics may represent the time required for the IMC to fold/bind the protease domain. Inhibition kinetics were analysed using two different approaches. First, concentrations of Sub^{WT} and Sub^{PC} were adjusted such that their velocities for cleaving *N*-AAPF-*p*-nitroanilide were identical. Suitable enzyme concentrations were selected^{22,25} and reactions were initiated by addition of Sub^{WT} or Sub^{PC}. IMC^{WT} inhibited Sub^{WT} more than Sub^{PC} (Fig. 2a), whereas IMC^{PC} initially inhibited Sub^{PC} more than Sub^{WT} but was subsequently degraded rapidly to release inhibition (Fig. 2b). A similar phenomenon can be observed with IMC^{WT}, but not in the timescale of these experiments. In the second approach, concentrations of Sub^{WT} and Sub^{PC} were identical and reactions were performed as before. IMC^{PC} bound weakly to Sub^{WT} when compared with IMC^{WT} (Fig. 3a), whereas IMC^{PC} bound Sub^{PC} initially with greater affinity, but was degraded more rapidly than IMC^{WT} (Fig. 3b).

These results indicate that IMCs imprint unique structural information onto their protease domain. Such structural imprinting, or 'protein memory', may be important for studies of protein folding. Our results also suggest that multiple, structurally distinct native structures can be encoded by unique amino-acid sequences. Although our studies were limited to subtilisin, this phenomenon addresses a general property of protein-folding mechanisms, as propeptides may not differ functionally from internal uncleaved sequences in proteins.

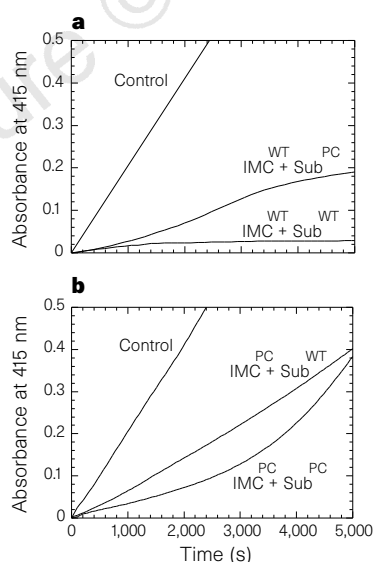


Figure 2 Comparison of inhibition profiles of Sub^{WT} and Sub^{PC}. Concentrations of the enzymes were maintained such that both Sub^{WT} and Sub^{PC} display the same units of activity towards the synthetic substrate *N*-succinyl-AAPF-*p*-nitroanilide. Reactions were performed at 23 °C, and substrate concentration maintained at 3 mM. Control depicts profiles of Sub^{WT} and Sub^{PC}, but only one has been shown because the two lines overlap. **a**, Inhibition profiles with IMC^{WT} (8.0 μ M); **b**, inhibition profiles for IMC^{PC} (3.74 μ M). IMC^{WT} interacts and binds tightly with Sub^{WT}, whereas IMC^{PC} binds tightly with Sub^{PC}.

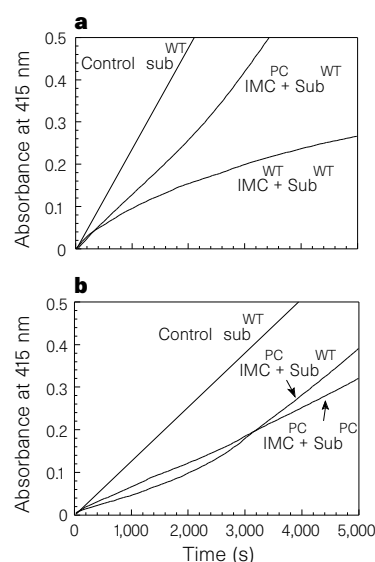


Figure 3 Comparison of inhibition profiles of Sub^{WT} and Sub^{PC}. The concentrations of Sub^{WT} and Sub^{PC} are identical (1.5 nM) but their units of enzymatic activity are different. Substrate and IMC concentrations were 3 mM and 1.0 μ M, respectively. **a**, Inhibition of Sub^{WT} by IMC^{WT} and IMC^{PC}; **b**, inhibition of Sub^{PC} by IMC^{WT} and IMC^{PC}. After 3,000 s, activity for the mixture of IMC^{PC} and Sub^{PC} is greater than that of IMC^{WT} and Sub^{PC}. This suggests that IMC^{PC} initially interacts and binds tightly to Sub^{PC} but is then rapidly degraded.

A phenomenon called enzyme memory was previously observed when subtilisin was lyophilized from aqueous solutions containing competitive inhibitors²⁸. Subtilisin then became much more active in anhydrous solvents than subtilisin lyophilized in the absence of the inhibitors. However, enzyme memory differs from protein memory in that enzyme memory disappears in aqueous solutions owing to the conformational flexibility of enzymes²⁹, whereas IMC-imprinted protein memory seems to be stable in aqueous environments. Another important implication of our finding relates to the maturation of eukaryotic hormones mediated by a family of subtilisin-like prohormone-convertases³⁰. Propeptides of some of these proteases function as IMCs, and it is possible that point mutations within the IMC domains may result in altered folding, and subsequent changes in substrate specificity of the enzyme may lead to hormonal disorders. □

Methods

Folding of subtilisin precursors. Refolding of precursors (50 µg ml⁻¹) was initiated through a dialysis procedure that involved stepwise removal of urea^{14,16,22}. Dialysis was continued for 48 h after complete removal of urea. The refolded proteins were incubated at 37°C for 6 h to facilitate propeptide degradation. This incubation can also be performed at 4°C for an additional 48 h without affecting the kinetic data of the proteins. The proteins were concentrated using an ultrafiltration membrane with a cutoff at an *M_r* of 10K, and were dialysed at 4°C for 24 h (using a tubing with a 12K cutoff) against refolding buffer (without urea) to remove contaminating peptide fragments that may have been generated by proteolysis²². They were subsequently purified using ion-exchange chromatography.

Slow-binding kinetics. The IMC domain of subtilisin behaves like a slow-binding inhibitor²². Experiments were therefore performed under pseudo-first-order kinetics, where the lowest IMC concentrations were greater than a 10-fold excess. Enzyme concentrations were set at levels low enough to give measurable rates of substrate hydrolysis, with observable state of inhibitor binding over the steady-state timescale. The IMC concentrations vary between 1 and 10 µM. Initial substrate concentration was 3 mM, and less than 10% of the substrate was degraded throughout the experiment. Reactions were initiated by addition of Sub^{WT} or Sub^{PC}. After thorough mixing, substrate cleavage was monitored by recording at 415 nm the release of *p*-nitroaniline. Data from each curve were fitted to the equation $A_{415} = v_s t + (v_0 - v_s)(1 - e^{-k't})/k' + A_0$, where A_{415} is the absorbance at 415 nm, A_0 is the initial absorbance, v_0 and v_s are initial and final velocities, respectively, k' is the apparent first-order rate constant, and t is time. Curve fitting gives four independent variables (v_0 , v_s , k' and A_0) at each inhibitor concentration. A plot of $(v_0 - v_s)/v_s$ with different inhibitor concentrations $[I]$ was used to determine K_i .

Circular dichroism. Circular-dichroism measurements were performed on an automated Aviv-60DS spectrophotometer controlled by an online temperature-controlled unit. Purified protein (100 µg ml⁻¹) in 10 mM phosphate buffer, pH 7.0, containing 0.5 M (NH₄)₂SO₄, 1 mM CaCl₂ was filtered through 0.22-µm filter and scans were carried out between wavelengths of 260 and 190 nm in a cuvette with a 1-mm path length maintained at 25°C. In the thermal unfolding experiments, temperatures were increased from 10 to 90°C in 1°C intervals, with a 30-s equilibration at each temperature. Data were collected at each temperature for 10 s. Protein solutions contain 1 mM PMSF to prevent autolysis during this procedure.

Received 21 February; accepted 3 July 1997.

1. Anfinsen, C. R. Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).
2. Kim, P. S. & Baldwin, R. L. Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* **59**, 631–660 (1990).
3. Ikemura, H., Takagi, H. & Inouye, M. Requirement of pro-sequence for the production of active subtilisin E in *Escherichia coli*. *J. Biol. Chem.* **262**, 7859–7864 (1988).
4. Ikemura, H. & Inouye, M. *In vitro* processing of pro-subtilisin produced in *Escherichia coli*. *J. Biol. Chem.* **263**, 12959–12963 (1988).
5. Winther, J. R. & Sorensen, P. Propeptide of carboxypeptidase Y provides a chaperone-like function as well as inhibition of the enzymatic activity. *Proc. Natl Acad. Sci. USA* **88**, 9330–9334 (1991).
6. Richo, G. R. & Conner, G. E. Structural requirements of procathepsin D activation and maturation. *J. Biol. Chem.* **269**, 14806–14812 (1994).
7. Shinde, U., Li, Y., Chatterjee, S. & Inouye, M. Folding pathway mediated by an intramolecular chaperone. *Proc. Natl Acad. Sci. USA* **90**, 6924–6928 (1993).
8. Zhu, X., Ohta, Y., Jordan, F. & Inouye, M. Pro-sequence of subtilisin can guide the folding of denatured subtilisin in an intermolecular process. *Nature* **339**, 483–484 (1989).

9. Eder, J., Rheinhecker, M. & Fersht, A. Folding of subtilisin BPN': role of the pro-sequence. *J. Mol. Biol.* **233**, 293–304 (1993).
10. Strausberg, S., Alexander, P., Wang, L., Schwarz, F. & Bryan, P. Catalysis of a protein folding reaction: thermodynamic and kinetic analysis of subtilisin BPN' interactions with its propeptide fragment. *Biochemistry* **32**, 8112–8119 (1993).
11. Silen, J. L. & Agard, D. A. The α-lytic protease pro-region does not require a physical linkage to activate the protease domain *in vivo*. *Nature* **341**, 362–364 (1989).
12. Lee, Y. C., Ohta, T. & Matsuzawa, H. A non-covalent NH₂-terminal pro-region aids the production of active aqualysin I (a thermophilic protease) without the COOH-terminal pro-sequence in *Escherichia coli*. *FEMS Microbiol. Lett.* **92**, 73–78 (1992).
13. Winther, J. R., Sorensen, P. & Kjelland-Brandt, M. C. Refolding of a carboxypeptidase Y folding intermediate *in vitro* by low-affinity binding of the proregion. *J. Biol. Chem.* **269**, 7006–7012 (1994).
14. Shinde, U. P. & Inouye, M. Folding mediated by an intramolecular chaperone: autoprocessing pathway of the precursor resolved via a 'substrate assisted catalysis' mechanism. *J. Mol. Biol.* **247**, 390–395 (1995).
15. Eder, J., Rheinhecker, M. & Fersht, A. Folding of subtilisin BPN': characterization of a folding intermediate. *Biochemistry* **32**, 18–26 (1993).
16. Shinde, U. P. & Inouye, M. Folding pathway mediated by an intramolecular chaperone: characterization of the structural changes in pro-subtilisin E coincident with autoprocessing. *J. Mol. Biol.* **252**, 25–30 (1995).
17. Baker, D., Sohl, J. L. & Agard, D. A. A protein-folding reaction under kinetic control. *Nature* **356**, 263–265 (1992).
18. Eder, J. & Fersht, A. R. Pro-sequence assisted protein folding. *Mol. Microbiol.* **16**, 609–614 (1995).
19. Shinde, U. P. & Inouye, M. Intramolecular chaperones and protein folding. *Trends Biochem. Sci.* **18**, 442–446 (1993).
20. Bryan, P. in *Intramolecular Chaperones and Protein Folding* (eds Shinde, U. & Inouye, M.) 85–98 (R. G. Landes, Austin, TX, 1995).
21. Kobayashi, T. & Inouye, M. Functional analysis of the intramolecular chaperone. Mutational hot spots in the subtilisin pro-peptide and a second-site suppressor mutation within the subtilisin molecule. *J. Mol. Biol.* **226**, 931–933 (1992).
22. Li, Y., Hu, Z., Jordan, F. & Inouye, M. Functional analysis of the propeptide of subtilisin E as an intramolecular chaperone for protein folding, purification and characterization of mutant propeptides. *J. Biol. Chem.* **270**, 25127–25132 (1995).
23. Fan, P. et al. Catalysis of a protein folding reaction: mechanistic implications of the 2.0 Å structure of the subtilisin-prodomain complex. *Biochemistry* **34**, 10310–10318 (1995).
24. Jackson, S. E. & Fersht, A. R. Contribution of residues in the reactive site loop of chymotrypsin inhibitor 2 to protein stability and activity. *Biochemistry* **32**, 13909–13916 (1993).
25. Longstaff, C., Campbell, A. & Fersht, A. R. Recombinant chymotrypsin inhibitor 2: expression, kinetics analysis of inhibition of chymotrypsin and wild-type and mutant subtilisin BPN' and protein engineering to investigate inhibitory specificity and mechanism. *Biochemistry* **29**, 7339–7347 (1990).
26. Mathews, B. W. Studies on protein stability with T4 lysozyme. *Adv. Protein Chem.* **462**, 240–311 (1995).
27. Russel, A. J. & Fersht, A. R. Rational modification of enzyme catalysis by engineering surface charge. *Nature* **328**, 496–500 (1987).
28. Russell, A. J. & Klibanov, A. M. Inhibitor induced enzyme activation in organic solvents. *J. Biol. Chem.* **262**, 11624–11626 (1988).
29. Klibanov, A. M. What is remembered and why? *Nature* **374**, 596 (1995).
30. Barr, P. Mammalian subtilisins: the long-sought dibasic processing endoproteases. *Cell* **66**, 1–3 (1991).

Acknowledgements. We thank B. Chait and U. Mirza for mass spectroscopy, and K. Madura, K. Chada, H. Berman and B. Brodsky for discussions and for reading this manuscript.

Correspondence and requests for materials should be addressed to M.I. (e-mail: inouye@rwja.umdj.edu).

Two-dimensional structure of plant photosystem II at 8-Å resolution

Kyong-Hi Rhee^{*†}, Ed P. Morris[‡], Daniela Zheleva[‡], Ben Hankamer[‡], Werner Kühlbrandt^{*†} & James Barber[‡]

^{*} European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, Germany

[‡] Wolfson Laboratories, Department of Biochemistry, Imperial College of Science, Technology and Medicine, London SW7 2AY, UK

The photosystem II complex, which is the most abundant membrane protein in chloroplasts, comprises the light-harvesting complex II and a reaction-centre core. The reaction centre uses the solar energy collected by the light-harvesting complex II to withdraw electrons from water, releasing oxygen into the atmosphere. It thus generates an electrochemical potential, providing the energy for carbon dioxide fixation and the synthesis of organic molecules, which make up the bulk of the biosphere¹. The structure of the light-harvesting complex II has been determined

[†] Present address: Max-Planck-Institut für Biophysik, Heinrich-Hoffmann-Strasse 7, D-60528, Frankfurt, Germany.

at 3.4-Å resolution by electron crystallography², but the high-resolution structure of the photosystem II reaction centre and other core components remained unknown. We have grown well-ordered two-dimensional crystals of a sub-core complex containing the reaction centre from spinach thylakoid membranes and used electron crystallography to obtain a projection map of its structure at 8-Å resolution. The features reveal the likely location of the key components that are active in electron transport, and suggest a structural homology and evolutionary links, not only with the purple bacterial reaction centre but also with the reaction centre of photosystem I.

The reaction centre of photosystem II (PSII) contains the D1 and D2 proteins, which bind the chlorophylls and pheophytins responsible for electron transport across the membrane. The reaction centre also contains the α and β subunits of cytochrome *b*-559 and the *psbI* gene product^{3,4}. Two chlorophyll *a* binding proteins (CP47 and CP43) link the reaction centre to the light-harvesting complex-like proteins CP29, CP26 and CP24, and to the outer antenna consisting of light-harvesting complex II (LHCII) (ref. 5). Three extrinsic proteins and a cluster of four Mn atoms on the luminal side of the PSII core form the oxygen-evolving complex (OEC). The sub-core we have crystallized consisted of the reaction centre and

CP47 (ref. 6), and therefore had lost CP43 and the OEC proteins. After detergent dialysis, the two-dimensional crystals obtained were tubular or sheet-like, measuring up to $3\ \mu\text{m} \times 2\ \mu\text{m}$. Cryo-electron micrographs of tubular crystals supported on thin carbon films were recorded. The two superimposed lattices of the flattened tubes were separated by image processing. Usually, both lattices were equally well ordered, and could therefore be used for structure determination.

Combined data from six lattices of four images are shown in Table 1. The lattice parameters of this crystal form are $a = 165.9\ \text{\AA}$ and $b = 162.9\ \text{\AA}$, with a 90-deg unit cell angle. The computer-generated diffraction pattern of an image is shown in Fig. 1. Data up to 8-Å resolution were used to calculate a two-dimensional map of the complex projected onto the membrane plane. The comparison of phases of symmetry-related reflections indicates that the crystals belong to the two-sided plane group $P22_12_1$. Data to 8-Å resolution had an overall phase residual of 19.7 deg (random = 45 deg), comparing a total of 1,883 independent phase measurements. Phase residuals in resolution ranges are given in Table 1.

The projection map (Fig. 2) shows that one unit cell contains four monomer PSII complexes, which are symmetrically arranged around two 2-fold axes perpendicular to the membrane plane and

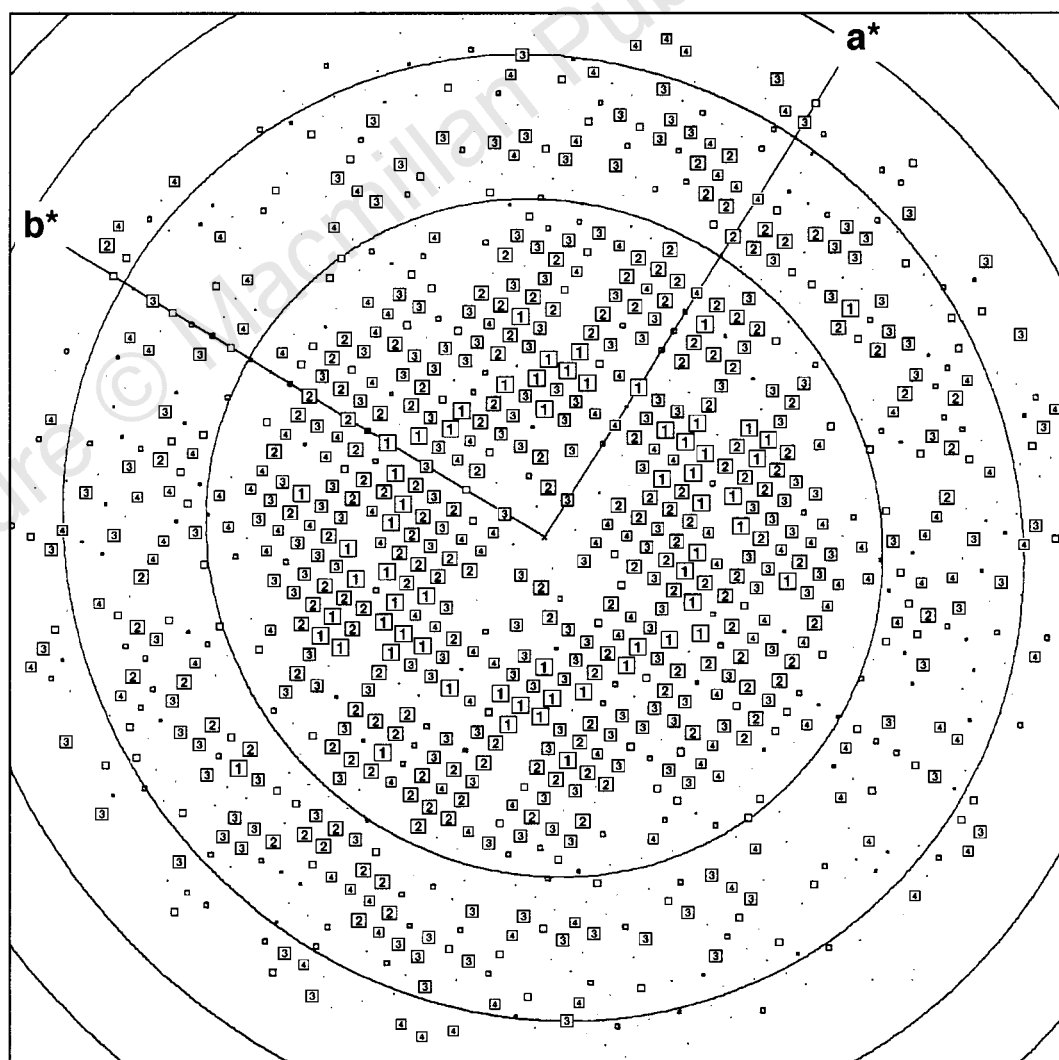


Figure 1 Fourier transform of an untitled image of a two-dimensional crystal of PSII. Each square on the reciprocal lattice indicates a Fourier component, of a size corresponding to the signal-to-noise ratio of the Fourier component¹⁴. The most significant structure factors are labelled with their IQ values, 1-4. Thin rings

indicating the zero crossings of the contrast transfer function are shown by concentric ellipses at resolutions of approximately 10.7, 7.6 and 6.1 Å. a^* and b^* refer to the reciprocal lattice vectors.

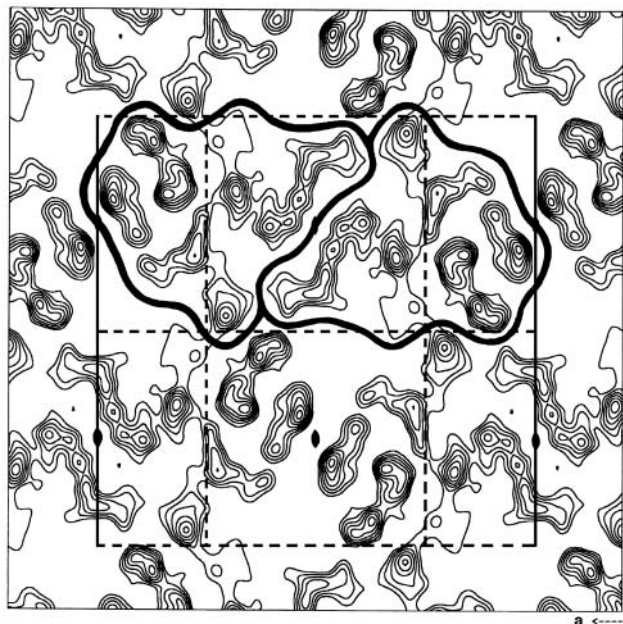


Figure 2 Projection map of PSII from spinach at 8-Å resolution, calculated from merged amplitudes and phases. One rectangular unit cell with $a = 165.9$ Å (horizontal) and $b = 162.9$ Å (vertical) is shown, with the symbols representing the crystallographic symmetry elements. There are two 2-fold axes perpendicular to the membrane plane and two 2-fold screw axes along both a and b axes. The unit cell contains four monomers. Two monomers, related by a crystallographic 2-fold axis, are outlined in bold. The outline of the monomer is known from low-resolution maps and from projection maps of crystals with different lattice parameters (not shown).

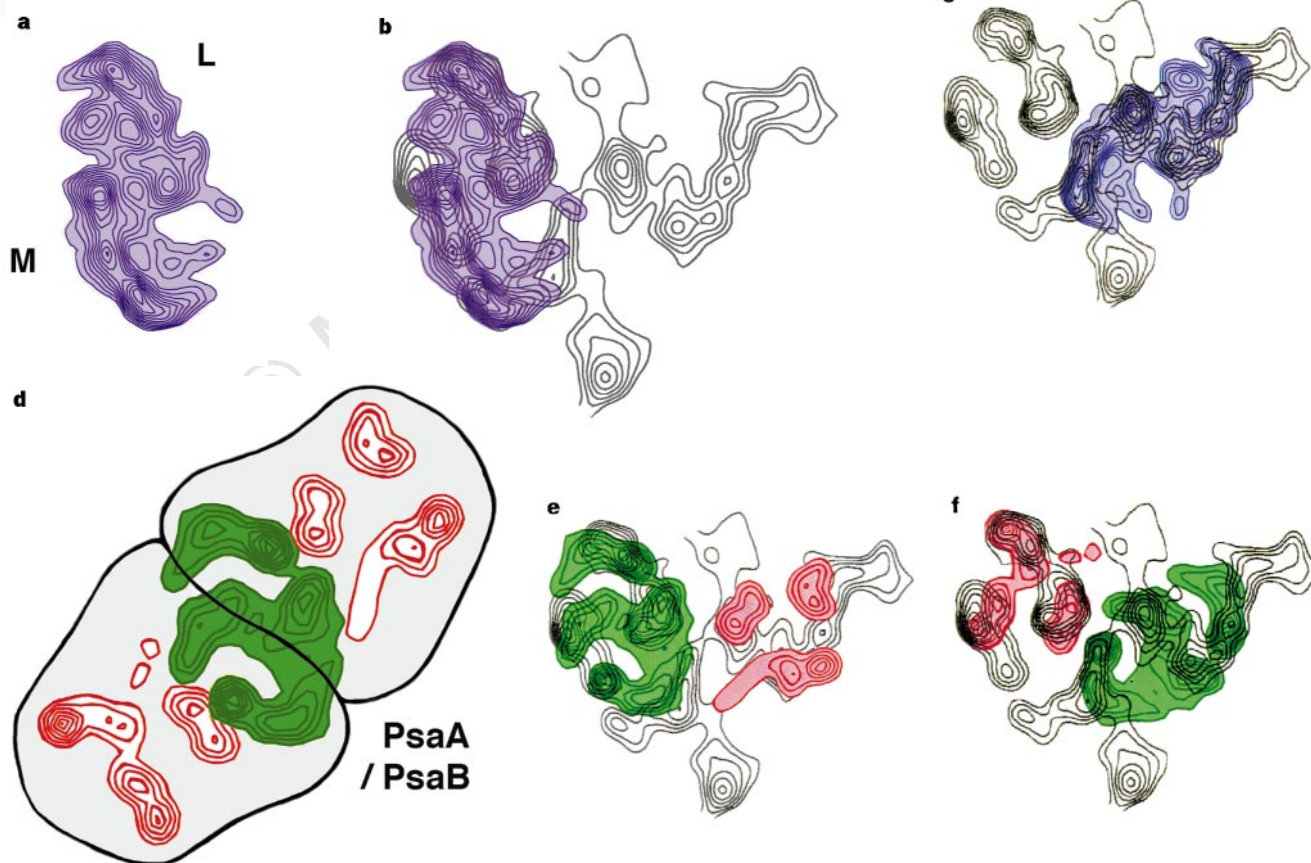


Figure 3 Comparison of the two-dimensional projected structures of the bacterial reaction centre, cyanobacterial PSI and plant PSII at 8-Å resolution. **a**, Projection map of the L and M subunits of the bacterial reaction centre from *Rhodospseudomonas viridis*. **b**, Superposition of the L and M projection map on the PSII sub-core complex of Fig. 2. **c**, Alternative positioning of the L and M projection map on the PSII sub-core complex. **d**, Projection map of the PsaA/PsaB subunit of PSI from *Synechococcus elongatus*. **e**, Superposition of the postulated D1/D2 projection map with the PSI inner-core reaction centre (green) and CP47 with the corresponding antenna region of PSI (red) generated by clockwise rotation of the PsaA/PsaB projection map in **d** around the pseudo-2-fold symmetry axis by 33°. **f**, Alternative positioning of the D1/D2 projection map

with the PSI inner-core reaction centre (green) and CP47 with the corresponding antenna region of PSI (red) generated by anticlockwise rotation by 15° to be completed. The bacterial reaction centre L and M subunits are coloured purple, PSII is black, the inner-core reaction centre of PSI (C-terminal five transmembrane α -helices of both PsaA and PsaB) is green, and the antenna regions (N-terminal six transmembrane α -helices) of both PsaA and PsaB are red. L, light and M, medium subunit of the bacterial reaction centre; D1, diffuse band 1 (*psbA* gene product); D2, diffuse band 2 (*psbD* gene product); CP47, chlorophyll binding protein of *M*, 47K; Cyt b559, cytochrome *b*-559; PSII-I, *psbI* gene product; PsaA and PsaB, PSI reaction centre core subunits.

Table 1 Electron crystallographic image data

Two-dimensional crystals				
Unit cell dimension			$a = 165.9 \pm 2.1 \text{ \AA}$ $b = 162.9 \pm 1.8 \text{ \AA}$	
Crystallographic space group			$P2_2,2_1$	
Density of protein in membrane plane ($\text{Da \AA}^{-2}$)			22.6	
.....				
Phase determination by image processing				
No. of processed images			4	
No. of merged lattices			6	
Resolution ($\text{\AA}$)			8	
Range of underfocus ($\text{\AA}$)			4,600–8,200	
Total no. of observed reflections			1,883	
.....				
Resolution range ($\text{\AA}$)	Averaged reflections*	Possible reflections	Completeness (%)	Phase residual (random = 45°)
in asymmetric unit				
200–15.8	88	95	93	9.5
15.8–11.2	77	88	88	19.0
11.2–9.1	78	91	86	25.7
9.1–8.0	64	77	83	27.6
Total 200–8.0	307	351	87	19.7

*Two or more observations of $IQ \leq 7$.

along two 2-fold screw axes in the a and b directions. The precise subunit composition of the cell is unknown, but it will contain at least the six components listed in Table 2, which total 19 transmembrane helices, assuming that all occur with a stoichiometry of one.

The amino-acid sequence of the D1 and D2 subunits of the PSII reaction centre suggests that these proteins contain five transmembrane helices each, and show significant homology to the L and M subunits of purple bacterial reaction centre⁷. A projected density map was generated from the α -carbon backbone of the L and M subunits of *Rhodospseudomonas viridis* (Fig. 3a). This map, viewed from the cytoplasmic side (equivalent to the stromal side in chloroplasts), shows similarity to part of the projection structure of PSII (Fig. 3b, c). This similarity was confirmed quantitatively by cross-correlation of the two maps. The alignment shown as an overlay in Fig. 3b gave rise to the highest correlation coefficient, suggesting that this is a likely position for the D1 and D2 proteins. A weaker correlation peak was found, however, when the map of the L and M subunits was aligned with another region shown in Fig. 3c. We thus consider this position for the D1 and D2 proteins as a possible alternative. The correspondence between the map of the purple bacterial reaction centre viewed from the cytoplasmic side and the map of Fig. 2 establishes the orientation of the latter with respect to the sidedness of the membrane.

A recent 4- $\text{\AA}$ structure of photosystem I (PSI) (ref. 8) indicates that the two PSI reaction-centre subunits, the PsaA and PsaB proteins (each composed of 11 transmembrane α -helices), have an inner-core reaction centre of two sets of five α -helices which are arranged very much like the transmembrane α -helices of the L and M subunits of the bacterial reaction centre. The inner-core region of

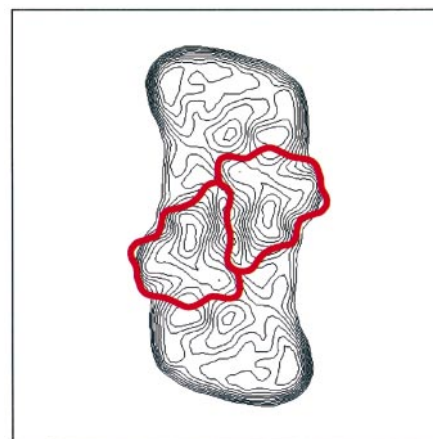


Figure 4 Superposition of the outline of the 8- $\text{\AA}$ projection map of the sub-core complex (red) with the PSII supercore complex dimer, as determined by single particle averaging¹². The supercore dimer map is viewed from the stromal side to match the PSII map outline. Thus the supercore projection map is mirrored with respect to the luminal view shown in ref. 12. The regions outside the red line contain the remaining PSII subunits, including CP43, CP24, CP26, CP29 and LHC-II.

a projected density map generated from the PSI data (Fig. 3d, green) can be aligned to produce quite good agreement with two possible regions assigned to D1/D2 in the PSII projection map by the overlay of the projection of the L and M subunits. The remarkable structural homology between the bacterial reaction centre and PSI thus also seems to extend to the PSII reaction centre. As can be seen in Fig. 3d, the remaining six transmembrane α -helices of both PsaA and PsaB form antenna regions composed of three pairs of α -helices that fan out on either side of the inner core. Some of these helices have amino-acid sequences that show significant homologies to the predicted six transmembrane helices of CP47 and may be evolutionarily important^{9–11}. We have therefore overlaid the corresponding regions of the PSI projection map (coloured red in Fig. 3d) with the major densities in the PSII map, which are not attributed to the D1 and D2 proteins in the two alternative positionings (Fig. 3b, c). These comparisons are shown in Fig. 3e, f. In both cases the best overlays required some realignment relative to the PSI structure and the regions identified as the D1 and D2 proteins. Nevertheless, both alignments look reasonable, with that shown in Fig. 3f being a better fit and requiring less realignment than that in Fig. 3e. To decide between the two possibilities will require further structural analyses, but the overlays presented in Fig. 3 indicate that the two photosystems active in plant photosynthesis, PSII and PSI, are likely to share structural similarities both in their reaction centres and in their inner antennae.

The PSII sub-core complex we obtained does not contain CP43, and so analysis of the crystals provides no direct information regarding its location. However, by analogy with the PSI structure, and bearing in mind the close homology between CP43 and CP47, it

Table 2 Known components of photosystem II present in the two-dimensional crystals

Gene	Location	Protein	M_r (K)*	Predicted transmembrane helices	Function
<i>psb A</i>	chloroplast	D1	38.8 (32)	5	Charge separation, electron transport, Mn-binding site, reaction centre
<i>psb D</i>	chloroplast	D2	39.4 (32)	5	Charge separation, electron transport, Mn-binding site, reaction centre
<i>psb B</i>	chloroplast	CP47	56.3 (47)	6	Chlorophyll a -binding inner antenna protein, excitation energy transfer
<i>psb E</i>	chloroplast	Cytochrome b -559 α	9.3 (9)	1	Haem-binding reaction centre protein, photoprotection
<i>psb F</i>	chloroplast	Cytochrome b -559 β	4.5 (4)	1	Haem-binding reaction centre protein
<i>psb I</i>	chloroplast	PSII-I	4.2 (4.8)	1	Unknown
			Total M_r , 153K		

*Relative molecular mass calculated from coding sequence. Value in parentheses is the apparent molecular mass.

seems likely that CP43 is located in an arrangement analogous to that proposed for CP47, but on the opposite side of the reaction centre (so, CP47 and CP43 would be approximately related by a 2-fold rotational axis passing through the centre of the D1/D2 region). Our projection map can be compared directly with the structure of the PSII supercore complex. This complex has been analysed by single-particle averaging^{9–12}, and is composed of a core dimer together with LHC and LHC-like proteins. The best agreement is found when using the dimeric complex from the PSII map outlined in Fig. 2, as this fits well into the region of the PSII supercore complex that is thought to contain the reaction centre and CP47 (Fig. 4). □

Methods

Two-dimensional crystallization. PSII was isolated from grana of spinach thylakoid membranes solubilized with *n*-heptyl- β -D-thiogluconide and crystallized as described in ref. 6 but with a slight modification. The chlorophyll concentration of the PSII OECs used for crystallization was 1 mg ml⁻¹ and the 100- μ l samples were dialysed against buffer containing 30% glycerol. When detergent was removed by dialysis the PSII complex forms well-ordered, large tubular crystals up to 3 μ m long after 1–3 weeks. The crystals are stable at 4°C for several weeks.

Electron cryo-microscopy. Specimens for electron cryo-microscopy were prepared on carbon film supported by copper grids in the presence of 0.5% tannin solution¹³ in dim light. Specimens were frozen rapidly in liquid nitrogen and mounted in a Gatan cryo specimen holder precooled with liquid nitrogen. Images were recorded at -182°C using a Philips CM200 field emission gun electron microscope equipped with a cold fork anticontaminator at an electron dose of 5–10 e⁻ Å⁻². Micrographs were recorded on Kodak SO-163 electron emulsion film with an acceleration voltage of 200 kV, a magnification of $\times 50,000$ and 1-s exposure time. Films were developed for 12 min in full-strength Kodak D-19 developer.

Image processing. Well-ordered crystalline areas of 6,000 $\times$ 6,000 pixels of the best four images were selected by optical diffraction and digitized on a Perkin-Elmer 1010 GM microdensitometer with a circular 10- μ m pixel at 7- μ m step size, corresponding to 1.4 Å at the specimen. Images were processed as described^{14,15} using MRC image-processing programs. Distortions of the two-dimensional crystal lattice were corrected by the program CCUNBEND, using a 400 $\times$ 400 pixel reference area in the first pass and 200 $\times$ 200 pixels in the final correction pass. As a result of the unbending procedure, the intensities of reflections increased on average by roughly a factor of two compared with raw data. After the correction of contrast transfer and objective-lens astigmatism, Fourier components from six lattices on four images with IQ value (a numerical grade of the signal-to-noise ratio of spots) ≤ 7 were merged by phase origin refinement. The maximum resolution was examined with the program HALFSTAT (provided by R. Henderson, MRC, Cambridge).

Two-dimensional map comparisons. To prepare Fig. 3a the α -carbon backbone from L and M subunits of *R. viridis* reaction centre was extracted from the Brookhaven protein database PDB file. Two-dimensional projection maps of the structure viewed from the cytoplasmic side were calculated. The resulting map was restricted to a resolution of 8 Å by band-pass filtering of the Fourier transform. Figure 3d was prepared in the same way except that the starting point was a set of three-dimensional coordinates defining the ends of the α -helices identified in a 4-Å map derived from X-ray crystallography⁸. Polyalanine α -helices were constructed between the helix ends and the resulting PDB file was processed as for Fig. 3a. The simulated projection maps were quantitatively aligned by cross-correlation against the experimental PSII data using the image processing system SPIDER¹⁶. The radial distribution of amplitudes in the Fourier transforms of the simulated projection maps was

adjusted to match that of the experimental data. This procedure improved both the discrimination between trial alignments and the absolute values of the correlation coefficients for the best fits.

Received 24 March; accepted 17 June 1997.

1. Barber, J. & Andersson, B. Revealing the blueprint of photosynthesis. *Nature* **370**, 31–34 (1994).
2. Kühlbrandt, W., Wang, D. N. & Fujiyoshi, Y. Atomic model of plant light-harvesting complex by electron crystallography. *Nature* **367**, 614–621 (1994).
3. Nanba, O. & Satoh, K. Isolation of a photosystem II reaction center consisting of D-1 and D-2 polypeptides and cytochrome b-559. *Proc. Natl Acad. Sci. USA* **84**, 109–112 (1987).
4. Barber, J. *et al.* Characterisation of a PS II reaction centre isolated from the chloroplasts of *Pisum sativum*. *FEBS Lett.* **220**, 67–73 (1987).
5. Jansson, S. The light-harvesting chlorophyll *a/b*-binding proteins. *Biochim. Biophys. Acta* **1184**, 1–19 (1994).
6. Nakazato, K. *et al.* Two-dimensional crystallization and cryo-electron microscopy of photosystem II. *J. Mol. Biol.* **257**, 225–232 (1996).
7. Michel, H. *et al.* The 'light' and 'medium' subunits of the photosynthetic reaction centre from *Rhodospseudomonas viridis*: isolation of the genes, nucleotide and amino acid sequence. *EMBO J.* **5**, 1149–1158 (1986).
8. Krauss, N. *et al.* Photosystem I at 4 Å resolution represents the first structural model of a joint photosynthetic reaction centre and core antenna system. *Nature Struct. Biol.* **3**, 965–973 (1996).
9. Vermaas, W. F. J. Evolution of heliobacteria: implications for photosynthetic reaction center complexes. *Photosynth. Res.* **41**, 285–294 (1994).
10. Rutherford, A. W. & Nitschke, W. in *Origin and Evolution of Biological Energy Conversion*, (ed. Baltscheffsky, H.) 143–174 (VCH, New York, 1996).
11. Fromme, P. *et al.* Structure of Photosystem I at 4.5 Å resolution: A short review including evolutionary aspects. *Biochim. Biophys. Acta* **1275**, 76–83 (1996).
12. Boekema, E. J. *et al.* Supramolecular structure of the photosystem II complex from green plants and cyanobacteria. *Proc. Natl Acad. Sci. USA* **92**, 175–179 (1995).
13. Wang, D. N. & Kühlbrandt, W. High-resolution electron crystallography of light-harvesting chlorophyll *a/b*-protein complex in three different media. *J. Mol. Biol.* **217**, 691–699 (1991).
14. Henderson, R., Baldwin, J. M., Downing, K. H., Lepault, J. & Zemlin, F. Structure of purple membrane from *Halobacterium halobium*: recording, measurement and evaluation of electron micrographs at 3.5 Å resolution. *Ultramicroscopy* **19**, 147–178 (1986).
15. Henderson, R. *et al.* Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* **213**, 899–929 (1990).
16. Frank, J., Shimkin, B. & Dowse, H. SPIDER-A modular software system for electron image processing. *Ultramicroscopy* **6**, 343–358 (1981).

Acknowledgements. We thank G. Friso for her involvement in the project, W. Schubert for supplying PSI data, and E. Orlova for advice on alignment by cross-correlation. This work was supported by the Biotechnology and Biological Sciences Research Council (J.B.) and by a fellowship to K.-H.R. from the Boehringer Ingelheim Fonds.

Correspondence and requests for materials should be addressed to W.K. (e-mail: kuehlbrandt@biophys.mpg.de).

addendum

Unaltered susceptibility to BSE in transgenic mice expressing human prion protein

John Collinge, Mark S. Palmer, Katie C. L. Sidle, Andrew F. Hill, Ian Gowland, Julie Meads, Emmanuel Asante, Ray Bradley, Lawrence J. Doey & Peter L. Lantos

Nature **378**, 779–783 (1995)

In this Letter we reported that transgenic mice expressing both human and mouse prion protein are no more susceptible to BSE than their wild-type counterparts. We also included the interim results of challenge of mice expressing only human prion protein (HuPrP^{+/+}Prn-p^{0/0}) with BSE; these mice had remained disease-free 264 days post-inoculation. We have now found that from 489 days onwards, some of these HuPrP^{+/+}Prn-p^{0/0} mice developed a prion disease. More details appear on pages 448–450 of this issue. □